Journal of
Experimental Psychology

VOL. 24, NO. 5 MAY, 1939

PATTERNS OF MUSCULAR ACTIVITY DURING 'MENTAL WORK' AND THEIR CONSTANCY

BY R. C. DAVIS

Indiana University

PROBLEM

With the accumulation of much evidence for the participation of skeletal muscles in psychological processes such as 'mental work' there remain to be answered a number of questions regarding the pattern of such activities over the muscular system and the relation of these patterns to the psychological process. Such questions have been the subject of much investigation when applied to neural functions; there would seem to be an equal importance to them as applied to the peripheral mechanisms. Three such problems with which this report is concerned are (1) What are the spatial patterns of muscular activity commonly occurring during psychological processes? (2) Are differing patterns characteristic of differing psychological processes? and (3) If muscle groups ordinarily active during a certain process are prevented from becoming so, will their function be taken over by other muscle groups?

In a few reported studies comparisons have been made between muscular activities found simultaneously in different parts of the body during some psychological activity, thereby giving information concerning the first of these three problems. In his well-known work on muscular phenomena during imaging Jacobson (**6**) recorded action potentials from both

arms while the subjects were imaging an activity of the right arm. He reported a highly localized activity in the right arm in his specially trained subjects. Shaw (7) found on normal subjects a wider spread of activity during similar tasks with muscular involvment decreasing with distance away from the member designated by the instructions. Such a distribution he found also characteristic to a certain extent of overt activities. In an experiment by the present writer (2) greater increase during problem solving was found to be characteristic of the forearm muscles as compared with the neck muscles. In a study of overt maze learning behavior Daniel (1) found muscular activities in three remote parts, their magnitude being some function of their distance from the active member. G. L. Freeman's study of the optimal locus of artificially induced tension (5) is rather closely related to the subject. The general result of his experiment was greater facilitation of the performance (finger oscillation) when the artificial tension was near the working finger.

Some suggestion of activity patterns characteristic of different psychological processes also comes from Jacobson (6) who presents evidence for arm activity during kinæsthetic imagery, tongue activity during verbal imagery and oculomotor during visual. Shaw's study likewise shows a different distribution of muscular activity during, for example, imaging weight lifting and imaging typing. The third problem, regarding shift or modification of muscular activity patterns during psychological activity, seems to be without experimentation thus far.

In the experiment here reported the objects in view are to compare muscular activity during mental work at three surface locations, to compare the spatial patterns of muscular activity during the two sorts of task, and to discover what changes in distribution occur when activity in one important muscle group is checked by voluntary inhibition.

Procedure and Apparatus

Two experimental groups of subjects (students) were used, the first group working at the multiplication of three place numbers by two place numbers, using the work limit method, the second, at memorizing lists of nonsense syllables, with the time

limit method. Although the use of two tasks somewhat hinders statistical work on most points, it seemed desirable to follow such a procedure, inasmuch as similar results obtained under somewhat different conditions would probably be more credible than results in a single situation substantiated by slightly higher critical ratios, and because the procedure makes possible a comparison of distribution for different tasks.

Subjects were seated in a chair, connected to the apparatus, and told to relax thoroughly for five minutes, to imagine they were going to sleep. The rest period was prolonged until a satisfactory rest level was reached (less than 5 microvolts, approximately), sometimes with the help of further, more specific instructions. Records were then taken. The subject was then instructed regarding the task, and the material to be worked on placed on a table before him. He was told to begin work when signalled to do so, and work just as hard as possible without changing position. (In learning nonsense syllables the 'whole method' was specified.) He was informed that he would be asked to give his answer or write the list of syllables at the end of the work period. Those working on multiplication were given just one problem at a time and told to signal the completion of the work. Those learning nonsense syllables were told to stop work at a signal which would come at the end of five minutes. (Actually they were allowed five minutes and fifteen seconds.) Half the subjects in each group were instructed to keep the right arm thoroughly relaxed during the first work period. In order to omit any shock effect, the first work records were taken fifteen seconds after the beginning of the work period. Subsequent records were taken at the end of each minute of work. The subject was then allowed to give his answer or write as many syllables as he could recall, regardless of order. No records were taken during this unspecified time. The second work period was then begun, with a second problem or list of syllables, the condition of attempted relaxation of the right arm being alternated. The signal to work was given only when the subject showed at least as good a level of rest as during the preliminary rest period. Two work periods were used for the nonsense syllable task, and four for the mental multiplication. Records were taken in the same manner as during the first work period. At the conclusion of the work periods a final rest period was given, during which records were taken. Before dismissed each subject was queried regarding his success in following the instructions and the difficulty encountered in working under the several conditions of the experiment.

Action potentials were used for recording muscular activity. The apparatus was essentially that described previously (3, 4), being two similar four stage amplifiers with gain controls feeding into cathode ray tubes, one three inches, the other five inches in diameter. Deflections on these tubes were recorded simultaneously on 35 mm paper in a moving film camera, along with the flashes of a gas glow tube connected to 60 cycle current. It was necessary from time to time during the work periods to reduce the gain of the amplifiers, particularly with certain subjects, measurements then being adjusted according to the calibration of the amplifiers. The usual settings of the gain controls gave sensitivities of .3 to .6 microvolt (A.C.) per 1/50 in (the unit in which measurements were made) on the recording paper, or about four times this amount before the reduction of the optical system.

The inactive electrode consisted of a saline soaked felt pad about 2 in $\times$ 5 in and $\frac{1}{2}$ in thick, placed on the upper right arm. Active electrodes were screw heads $\frac{1}{4}$ in in diameter moistened with saline and held in place with elastic bandages. As nearly as possible corresponding anatomical points were used on different subjects by placing electrodes on the highest crest on the back of each forearm and similarly on the calf of the left leg. The right arm electrode was connected to the one recording system at

all times; the two electrodes from the left side were connected alternately by means of a manually operated commutator allowing a record from one position to follow that from the other by a few seconds. Since each run of the camera phtoographed both recording systems there were twice as many records from the right arm as from either of the other two locations. The camera runs lasted approximately 3–5 seconds.

The method of measurement and treatment of data was substantially the same as that used in an earlier investigation (2). From each camera run three sample periods, $\frac{1}{12}$ second in duration, were selected, one near the beginning, one in the middle, one near the end, care being taken to avoid the electrocardiogram which inevitably appears on the records from the left side under these conditions. In each of these periods the difference between the most positive and the most negative point was measured in 1/50's of an inch, for both recorded lines. The three measurements for each electrode location in each camera run were then averaged and converted to microvolts. For the main purpose of the investigation further averages were then computed for the whole rest periods and the whole time spent in work under the two conditions of the experiment. As in the preceding investigation, averages for the two conditions of work were then expressed as percents of the average of both rest periods. For the present purposes the percentage technique seems preferable to a difference technique since it would be expected to eliminate gross individual differences produced by variation in location of the electrode with respect to the acting muscle, and other such accidental factors.

Results

(1) *Reliability of Measurements.*—Any technique of measuring and quantitatively treating records of muscle action potentials is somewhat arbitrary, and it is desirable to have some indication of the statistical reliability of the one chosen. Two sorts of reliability have been computed from the present data. Two subjects showing the smallest range of readings were selected from the first experimental group. On their records new sample periods of 1/12 second each were marked off adjacent to the original samples during the work periods. These were measured by assistants having no knowledge of the original data. Rank difference correlations between the new and old measurements were found to be, for subject R, right arm: .83, left arm: .58, left leg: .67; for subject W, right arm: .94, left arm: .76, left leg: .95. Correlations would of course have been higher if the rest period measurement had been included, but in any case the figures are high enough to justify considerable reliance on the average reading for an individual. Such, however, is not the kind of average with which we are most concerned. To arrive at the reliability of a group average during the continuance of particular condi-

tions the split fraction technique was employed. Each camera run contains three independent measurements. The correlations for the first two measures for the first camera run of the first work period were computed for the subjects working on multiplication. The rank difference correlations for the right arm uncorrected by the prophecy formula are .89 for group A, .97 for group B. These are evidently of satisfactory magnitude, especially when it is considered that the averages used for making comparisons are generally of higher order than that of the single camera run.

Temporal Patterns.—It was noticed during recording the action potentials were not always equally distributed through the work period. An initial burst (which was not recorded) was usual, but subsided by the end of 15 seconds, so that the first reading is often little greater than the rest readings. Thereafter there was a tendency for muscular activity to increase, sometimes rather suddenly. To obtain some measure of the effect the last camera run in each work period was expressed as a percentage of the first. These percentages are shown in Table 1, and seem to suggest that the tendency is

TABLE 1

RATIO OF LAST WORK PERIOD READING TO FIRST WORK PERIOD READING, IN PERCENT

	L. Leg		L. Arm		R. Arm	
	Normal	'Relax'	Normal	'Relax'	Normal	'Relax'
Group A Arith.....	155	113	153	126	133	109
Group B Learning..	114	104	121	115	158	67

strongest during work under normal conditions, and especially prominent in the right arm, being probably absent altogether in that arm under instructions to relax it. Possibly the effect is associated with increasing difficulty of task as the work progresses.

Relation of Muscular Activity to Work Output.—Though not especially studied in this investigation such a relationship may readily be computed. The work measure for subjects of the

first group was the reciprocal of the time spent; for the second, the number of nonsense syllables correctly recalled. Work periods were combined without regard to relaxation instructions. The measure of muscular activity was the average figure for a subject's work period expressed as a percent of his rest level. The rank difference correlations are .48 for the multiplication, and −.12 for the nonsense syllables learning. These figures agree with an earlier conclusion (3) that the relation between muscular activity and output of work is not a simple one and may be represented by a correlation of any magnitude, depending on the disposition of other influential factors.

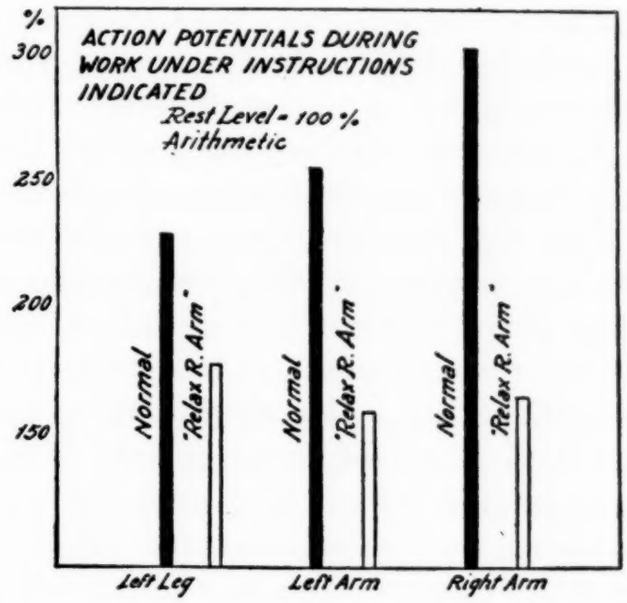
Spatial Patterns of Muscular Activity.—Averages of action potential readings at various points are given, with their statistical constants, in Table 2, for both groups and both

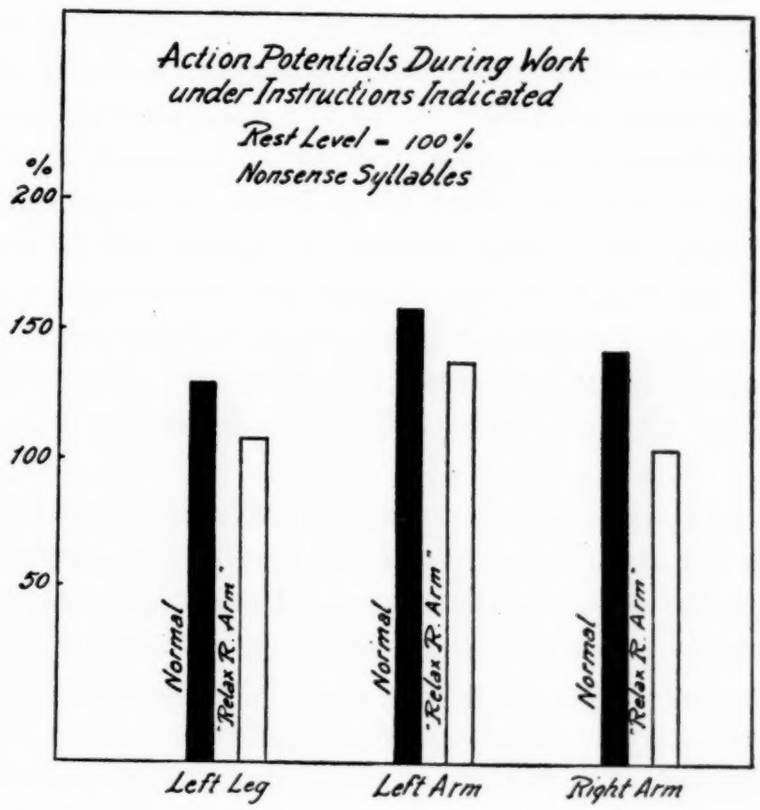
TABLE 2

COMPARISON OF ACTION POTENTIALS DURING NORMAL WORK WITH ACTION POTENTIALS DURING WORK WITH INSTRUCTIONS TO RELAX RIGHT ARM. FIGURES IN TERMS OF PERCENT OF REST LEVEL

Group Constants		L. Leg		L. Arm		R. Arm		Work Measure	
		Normal	'Relax'	Normal	'Relax'	Normal	'Relax'	Normal	'Relax'
A Multiplication	Av...	231.7	178.5	256.3	161.5	316.7	167.8	140.5″	125.5″
	S.D. .	173.6	127.8	163.1	47.8	220.5	52.6		
N = 15	ρ....	.68		.55		.76		.71	
	Diff..	53.2		94.8		133.8		15.0	
	$\sigma_{diff.}$	35		38		48		14.1	
	$\frac{D}{\sigma_{diff.}}$	1.2		2.5		2.9		1.1	
B Nonsense Syllables	Av...	133.5	110.1	160.3	142.9	145.5	107.6	8.63	7.25
	S.D. .	53.8	33.3	62.6	61.9	50.6	48.7	3.40	2.57
N = 16	ρ....	.55		.63		.33		.83	
	Diff..	23.4		27.4		37.9		1.38	
	$\sigma_{diff.}$	11.3		13.3		14.6		.48	
	$\frac{D}{\sigma_{diff.}}$	2.1		2.0		2.6		2.9	

sorts of work conditions. Figures 1 and 2 represent these averages for the two groups. The table as a whole is based on approximately 5000 measurements.





There seems to be little question but that there is a substantial increase in muscular activity in the locations studied during the performance of the tasks in question under the 'normal' conditions of work in both groups. In the first group 98 percent of the 'normal' work readings for different locations on individual subjects are greater than the same subjects' rest averages. In the second group the corresponding figure is 85 percent. During work under instructions to keep the right arm relaxed the readings for the first group are all greater than the rest averages. In the second group, however, the percentage exceeding rest averages is only 68, reversals being most frequent in the right arm. It will be recalled that the rest averages are based on readings before and after the work period. If work readings in this last group are compared with final rest readings only, it is found that 80 percent are greater. Thus it is probable that even here work level is higher than a genuine rest level.

The diagrams taken at their face value indicate a focus of muscular activity in certain places with a decreasing gradient away from these. The focus for the multiplication seems to be in the right arm, while in learning nonsense syllables neither arm dominates the other. Explanation for such a difference is suggested by occasional spontaneous reports from subjects that they had a strong tendency to write during the multiplication, a report which was never given by the subjects working on the other task. (All subjects in the multiplication group were right handed. One in the group learning nonsense syllables was left handed, but showed no characteristic difference from the rest of the group.)

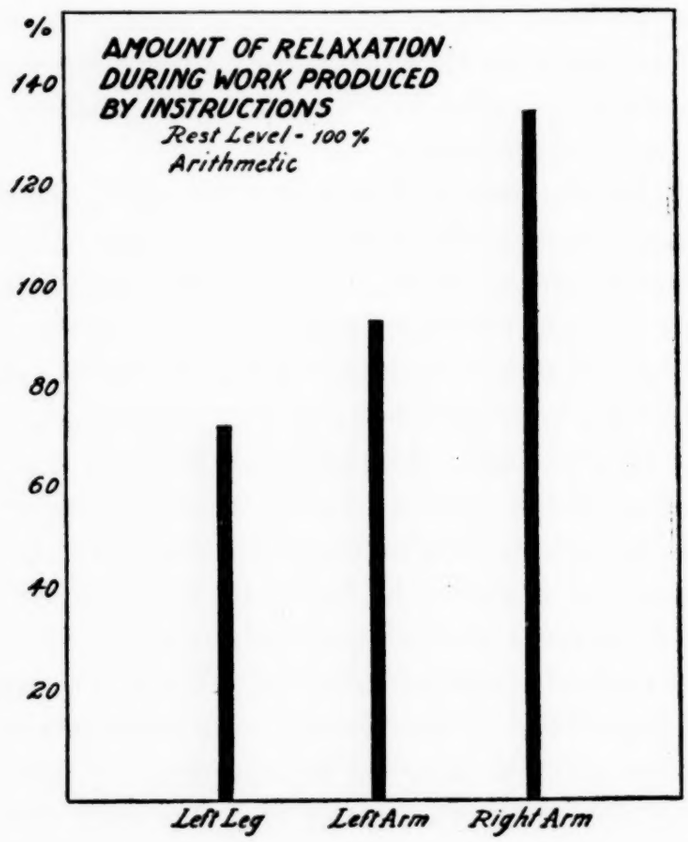
Yet the critical ratios of the differences between the several locations are low, being about 1.5 in all cases except for the two arms during multiplication, and one must guard against giving much weight to the figures except where one group confirms the other, in which case the reliability would be quite satisfactory. There is thus confirmation for a gradient of activity between the upper and lower limbs. The gradient from right to left for multiplication possibly exists, but needs further checking.

There is a conspicuous difference in the level of activity shown during the two sorts of tasks. Averaging the work level regardless of instructions to the subject at the three points (in percent of rest level) we have, for the multiplication, 214.7, for the nonsense syllable learning, 131.5, with a difference of 83.2, which is 3.1 times its standard error. Evidently there is a considerable difference in the stimulating power of the two tasks for the regions measured. Activity of the most active part during learning nonsense syllables is less than activity in the least active part during multiplication. It is possible of course that there is a highly concentrated activity during nonsense syllables learning in some unmeasured part, say the speech musculature.

Effect of Instructions to Relax One Member During Work.— In Table 2 are shown the differences and their critical ratios between work under normal conditions and work under instructions to keep the right arm relaxed. Comparison may also be made in the four figures. Subjects clearly did succeed in maintaining a lower level of activity in the right arm under the special instructions, though generally not succeeding in maintaining as low a level as during the rest periods.

There is no evidence whatever of a compensating increase in other parts to replace the tension of the right arm. On the contrary muscular activity elsewhere is likewise inhibited, and by a statistically reliable amount when the two groups are considered together. The effect of the instructions to keep the right arm relaxed may be represented by the difference between the normal work period and the work period under the special instructions. These differences are represented graphically in Figs. 3 and 4. The pattern of the effect is strikingly similar in the two groups. The inhibition of the right arm activity seems to extend to other regions in relation to their distance from the right arm. It may be suggested that we are dealing with a right arm task, as in the case of multiplication, and that inhibition follows the same rule of spread as excitation. Naturally differences in amounts of inhibition, being of the second order, are difficult to establish statistically. With the exception of one, which is so low as

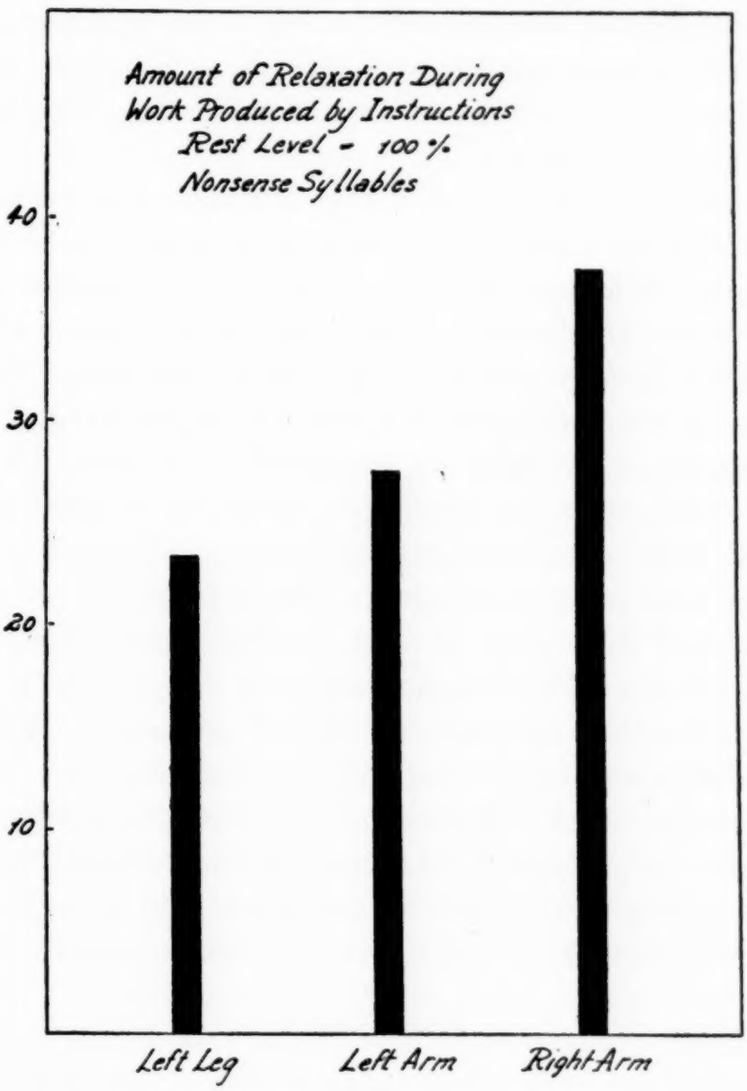
to be quite insignificant, critical ratios for the differences between amounts of inhibition lie between 1.0 and 1.5, enough to give a reasonable degree of assurance when the phenomenon occurs twice. Yet there is a question whether such a simple technique as the critical ratio is applicable in this case. The question here is whether the trend line of the data has a slope significantly different from zero, and the occurrence of the same trend twice is rather convincing evidence that it does.



It is reasonable to conclude that attempted relaxation of the right arm during mental work produces a reduction of tension in that arm but without abolishing tension, and a lesser reduction in the tension of other members in relation to their distance from the right arm.

Since the subjects were assigned the two tasks of voluntary relaxation and work it is pertinent to inquire whether these actually proceeded concomitantly, or whether there was any alternation between them. It is possible to answer the

question by reference to the muscular action potential data for the reason that the readings are systematically scattered through the duration of the work period. Now if a subject were to show alternately the tension level of 'normal' work and a low level of tension while he was momentarily trying to relax, the effect would be to increase the scatter of muscular



activity readings for him. The average of these readings could, of course, be lowered from normal either by such an alternation or by a consistent lowering of level such as would be obtained from actual concomitance of the two tasks. Hence, it is important to compare the standard deviations of the two series of readings for each individual, the one under

normal work conditions, the other under the special instructions. Data in terms of microvolts were used for this comparison. For the group working on multiplication the average of the individual standard deviations is 15.3 during normal work, 8.3 during work with the differential relaxation (figures for the right arm). For the second experimental group corresponding figures are 1.55 and 1.22. Evidently the scatter of readings is actually less rather than greater under the special instructions, and it is justifiable to conclude that work and differential relaxation were carried on simultaneously rather than alternately.

Effect of Relaxation on Work Accomplished.—One would expect the general pattern of relaxation which is as produced by the special instructions to reduce the output of work for two reasons. In the first place, since mild tension facilitates work (Bills) its opposite would probably interfere. Secondly, the subjects were of course assigned the double task of working and relaxing and might be expected to accomplish less on that account. For the first group, working on multiplication, only the time measures are available. (Only one subject arrived at any correct answers.) Reference to Table 2 shows no significant difference in time worked under the two sorts of instructions. The second task was so chosen as to yield a more complete measure of output of work. With time constant number of items correctly reproduced must represent the output of work. Here there is definite evidence for a decrease in the output, as shown in Table 2, while the subject was trying to keep the right arm relaxed (and actually inhibiting tension elsewhere also) during the work period.

Discussion

The results of the investigation may conveniently be described as resulting from the operation of certain principles.

1. The principle of muscular excitation during mental work. This principle is of course by no means new, having been found to operate in a great number of previous investigations. In the present study, it is followed in practically all subjects, a small minority showing slight reversals. These

reversals could well be accounted for by let-downs during work periods, and failures to secure good relaxation during rest periods.

2. The principle of focus of muscular responses. In the occurrence of any psychological process there is some region, constant from subject to subject, in which there is a relatively high degree of muscular activity. Such a principle is indicated by the studies of Shaw, Daniel and Davis previously referred to. In the tasks examined the maximum activity is located in the upper limbs in both cases, and for the multiplication activity there is some evidence that the center of still higher activity is located in some other position which was not studied. This principle of focus holds for inhibition of muscular activity as well as for excitation.

3. The principle of response gradient. Parts of the body show a lessened magnitude of response with increasing distance from the focal part. This principle also operates for inhibition as well as for excitation. This principle applies to the points measured in the present experiment and to the results of certain other experiments as previously cited, but thus far there its universality has not been demonstrated, since large intervening areas remain unstudied.

4. The principle of correlation of response. The excitation or inhibition of activities in various parts of the body tends to be concomitant. When one is excited, others are to some degree and when one is inhibited, others are, to some degree. This has been the trend of results in previous experiments, and is confirmed for inhibition by the present results, which show a concomitant variation in other parts during differential relaxation, rather than a compensation.

5. The principle of specificity of muscular action pattern to task. In the present experiment it may be said there are four different tasks assigned subjects: multiplication and learning nonsense syllables, working under normal conditions, or under special instructions regarding relaxation. For each of these tasks there is a characteristic sort of muscular activity, different in spatial pattern and intensity from the others. How far this principle of specificity is valid for other tasks, of

course, remains to be tested. Much depends on the fineness with which one classifies tasks. For example, if it is asserted that subjects are doing the same task while they are working under normal conditions and under relaxation instructions, it would seem that we have two patterns for the same task. However, it seems better to regard the two sorts of instruction as constituting two varieties of assigned task.

There may also be in operation a sixth principle, namely, the lowering of work output when there is interference with muscular activity. This, however, is more of a hypothesis for further testing by some technique which does not require the subject to perform a double task.

It is possible to relate the results of the present experiment to the account of the conditions affecting muscular activity during mental previously given by the writer (**2, 3**). In the earlier papers a functional equation was set up to describe the influencing factors. In the light of the present data two factors should be added to that equation: one factor a constant of excitability for the part of the body concerned, the other representing the location of focus and steepness of excitation gradient for any particular task.

Summary

1. Action potentials from three limbs were measured in two groups of subjects while they worked at mental multiplication and learning nonsense syllables under normal working conditions and under instructions calling for a relaxation of the right arm during work.

2. The reliability of action potential measurements by the method used was tested and found satisfactory.

3. Muscular activity was found to increase progressively during work periods.

4. No consistent relation between action potential and output of work was found under the experimental conditions when individual subjects were compared.

5. Subjects were found to show more muscular activity in all parts during mental work than during rest in almost all cases.

6. Muscular activity was found to be at a maximum consistently in one part and to be present in others in decreasing amount, varying with their distance from the focus.

7. It was found that instructions to relax the right arm during work produced a decrease of activity there with smaller decreases in the other parts, with no discoverable compensation.

8. The results were interpreted in terms of principles of excitation during mental work, focus, gradients and correlations, muscular excitations, and specificity of pattern to task set.

(Manuscript received November 26, 1938)

References

1. Daniel, R. S., *The Distribution of Muscular Action Potentials during Maze Learning*, to be published, this journal, June 1939.
2. Davis, R. C., The relation of muscle-action potentials to difficulty and frustration, *J. Exper. Psychol.*, 1938, **23**, 141–158.
3. Davis, R. C., The relation of certain muscle action potentials to 'mental work,' *Indiana Univ. Sci. Ser.*, 1937, No. 5, pp. 29.
4. Davis, R. C., The muscular tension reflex and two of its modifying conditions, *Indiana Univ. Sci. Ser.*, 1935, No. 3, pp. 27.
5. Freeman, G. L., The optimal locus of 'anticipatory tensions' in muscular work, *J. Exper. Psychol.*, 1937, **21**, 554–564.
6. Jacobson, E., Electrophysiology of mental activities, *Am. J. Psychol.*, 1932, **49**, 677–694.
7. Shaw, W. A., The distribution of muscular action potentials during imaging, *Psychol. Rec.*, 1938, **2**, 195–215.

EXPERIMENTAL STUDIES IN ROTE-LEARNING THEORY. IV. COMPARISION OF REMINISCENCE IN SERIAL AND PAIRED-ASSOCIATE LEARNING

BY CARL IVER HOVLAND

Department of Psychology, Institute of Human Relations, Yale University *

Introduction

Short-time reminiscence, which is very pronounced in serial learning with a two-second rate of syllable presentation, practically disappears when the time interval between successive syllables is increased (Hovland, **4**). This reduction in reminiscence might be attributed to a decrease in intra-serial interference as a result of the longer period of time between successive syllables. Certain interesting implications would then follow concerning reminiscence in paired-associate learning. In serial learning of the type in which reminiscence has been obtained, not only immediate but also remote associative bonds are formed. (Ebbinghaus **2**, McGeoch **11**, and others.) These remote bonds have been assumed to give rise to interference ('inhibitory') effects (Hull, **7**). In paired-associate learning, in contrast to serial learning, the associations are principally between the paired syllables and the establishment of remote associations is very infrequent. This is especially the case when one employs the procedure of presenting the paired-associates in a constantly changing order so that only the two members of the pairs remain in a fixed relationship. If the formation of remote associational bonds is an important factor in producing interference, and if reminiscence results from a decrease in interference effects with the

* This is the fourth article in a series of studies by the writer and collaborators on the experimental implications of a logico-mathematical system of serial learning to be published (**8**).

The writer wishes to thank Professors Clark L. Hull, John A. McGeoch and Donald G. Marquis for critical reading of the manuscript.

passage of time, reminiscence should then be much less prominent in paired-associate than in serial learning. Comparison of the amount of reminiscence with the two methods of learning is the problem of the present experiment.

That the difference in the experimental procedure employed in the serial and in the paired-associate learning methods greatly affects the phenomena obtained has been brought out forcibly in Kjerstad's (9) experiment. He found that paired-associate materials are learned at a relatively constant rate throughout the learning period, whereas serial lists are learned at a negatively accelerated rate. Only about 30 percent of the total amount of serial material is learned during the last half of the learning time, whereas with paired-associates the amount is nearer 50 percent. The percentages of the total material learned during successive thirds of the learning are 56, 28 and 16 percent with serial materials, but 35, 41 and 24 percent with paired-associates. It is possible that this difference is also a function of increased interference effects in serial learning, and hence is closely related to the phenomenon under investigation. A later article in the present series will be devoted to this problem.

Experimental Procedure

The experimental procedure was designed to determine the extent of reminiscence in paired-associate learning with exactly the same conditions under which it has been obtained with the serial anticipation method (3, 4). To equate the number of trials required for learning by the two methods, a preliminary experiment was run. The results showed that subjects require as many trials to learn a list of nine paired-associates as a serial list of eleven syllables, and ten pairs require more trials than twelve serial syllables. The paired-associate lists were, therefore, made up of nine paired nonsense syllables so as to permit the closest comparison with the results obtained previously on twelve-unit serial lists (4).

Reminiscence was studied by measuring the effect of a two-minute rest pause interpolated at two stages of learning: (1) after six syllables were correctly anticipated, and (2) following the first perfect recitation of the entire list. Results on a third stage, *i.e.*, following a single presentation of the list, were obtained from an experiment by the writer on the distribution of practice with paired-associate materials (6), in which all experimental conditions were identical with those of the present study. The second trial (first attempt at recall) with massed practice was compared with the second trial with distributed practice. In the former the recall was immediately after the first presentation of the list; in the latter, after a two-minute rest period of color-naming.

The three points of interpolated rest were identical with those used in the previous

experiment employing the serial anticipation method, except that with the serial method the rest pause at partial learning was introduced after learning to a criterion of seven out of twelve syllables correct, while with the paired-associate material the criterion was six out of nine correct.

The subjects learned the paired-associate materials in balanced practice orders with the following experimental programs:

I. *Learning to Criterion of One Perfect Recitation by Massed Practice.*

II. *Learning to Criterion of Six Syllables Correct by Massed Practice; Two-minute Rest Pause; Continuation of Learning by Massed Practice to One Perfect Recitation.*— On the first trial on which six or more syllables were correctly anticipated, the shutter in front of the window where the syllables were exposed was closed and a shutter exposing colors was opened. The subjects named these colors as presented. This activity effectively prevents rehearsal of the syllables and gives little retroactive inhibition. Following two minutes of color-naming this shutter was again closed, the syllable shutter reopened, and the subjects continued their learning.

To obtain data on the effect of the rest pause following mastery, the cases involving Program I were subdivided, half of them continuing to a second perfect recitation immediately, the others after a two-minute period of color-naming. Four sessions employing Program I were thus obtained, but only two of Program II were required.

The experimental set-up was identical with that used in the previous experiments (3, 4). The experiment was performed in a sound-proof chamber to eliminate distractions. The anticipation method of presenting paired-associates was employed to permit closest comparison with the results obtained with the serial anticipation method. The Hull exposure apparatus was used. This was arranged to give a two-second exposure of the first syllable followed by a two-second exposure of its associate. The list then moved on to the next pair, exposing each member successively for two seconds. This continued until all nine pairs had been presented. Following a six-second pause, the series was presented again with the order of the successive pairs altered. The lists were made up with six arrangements of the nine pairs in such a way that the same pairs never appeared consecutively on successive presentations of the lists. Six practice and six experimental lists were prepared. The standard rules for construction of nonsense syllable lists were followed. The lists were equated in difficulty according to the association frequencies of Hull. Practice orders were prepared in advance which were balanced with respect to all the significant variables.

During the six practice periods (days 1, 2, 3, 4, 5 and 6) the subjects learned two of the lists according to Program I with immediate relearning, and two with relearning after a two-minute rest pause. Two lists were also learned with Program II. The program used on a given day was arranged in advance to give counterbalanced practice orders. The data obtained on days 7, 8, 9, 10, 11 and 12 were used for the main experiment. The practice order for a given subject was never the same as the one used for him during the practice sessions.

The instructions to the subject were these:

"This is an experiment in learning nonsense syllables, and not a psychological test. We are interested in certain complex relationships of the learning process common to all people, and not concerned with your personal reactions.

"Shortly after the apparatus starts you will see a three-letter syllable in the first window (Indicated by the experimenter). After a few seconds this window will close and the adjacent one open. You are to pronounce each of these syllables as you see them. Together they represent a pair. You are to learn to associate the two so that

when the first appears you can give the second before it appears. The pairs themselves will not follow each other in any regular order but the two members of the pair will always occur together.

"Following the first complete exposure of the entire series of pairs you are to begin to anticipate the second syllables of the pairs; in other words, you are to pronounce the second syllable of the pair before it appears, while the first is still exposed. If you think you know what the syllable is, but are not sure, guess, because it will not hurt your score any more than to say nothing, and if you get it right it will count as a success. If you anticipate a syllable incorrectly, correct yourself as soon as it appears. Try always to speak the syllables as distinctly as possible.

"Please do not try to think ahead more than one step at a time, or to count, or to make up fanciful connections between the syllables to assist the learning process. Don't try to use any special system in your learning; simply learn each syllable with its associate. When both shutters go down, begin naming the colors as they are presented. Do not ask questions about the purpose of the study until the experiment is over."

The thirty-two subjects employed were Yale College students, who were paid for their services. They had not previously been employed in memory experiments and were not aware of the purpose of the experiment. The program to be used on a given day was not announced to the subject in advance.

Results

Reminiscence Following a Single Presentation.—Table I shows the relative retention with and without a rest pause following the first presentation of the lists in serial and in paired-associate learning. As mentioned above, these data were collected in an experiment by the writer (**6**) to determine the relative effectiveness of massed and distributed practice with the two methods of learning. In massed practice learning, the first recall follows the initial presentation immediately, whereas with distributed practice there is a two-minute period of color naming between presentation and recall. Thus, these results furnish data upon reminiscence following a two-minute rest pause. It will be observed from the table that with serial learning definite reminiscence effects are obtained. The recall following the rest pause is reliably higher than the recall on the comparable trial without a rest pause (C.R. = 4.32). This is in accord with results previously obtained by the writer (**3, 4**) and by Ward (**14**). When, however, the two-minute period is introduced following the first presentation in the paired-associate learning, there is not only no improvement ('reminiscence'), but instead an actual loss ('forgetting') with the passage of time (C.R. = 2.23). This difference

TABLE 1

RETENTION FOLLOWING ONE PRESENTATION OF NONSENSE SYLLABLES SERIES BY SERIAL AND PAIRED-ASSOCIATE METHODS WITH AND WITHOUT 2-MINUTE REST PAUSE BETWEEN PRESENTATION AND RECALL

(A) Recall Scores

(32 subjects, 2 determinations for each condition)

	Paired-Associate		Serial	
	Mean	σ_M	Mean	σ_M
No. syllables correctly recalled immediately after presentation	0.67	0.103	0.75	0.079
No. syllables correctly recalled 2 minutes after presentation	0.41	0.082	1.16	0.087
Difference	−0.26		+0.41	
$\sigma_{diff.}$*	0.116		0.095	
C.R.$_\sigma$	2.23		4.32	

(B) Cases and Subjects Showing Reminiscence

(64 paired cases)

	Per cent	σ_P	Per cent	σ_P
1. Paired trials in which more syllables were recalled following 2-minute period than immediately	14.1	0.044	40.6	0.062
2. Paired trials in which same number of syllables were recalled	53.1	0.062	32.8	0.059
3. Paired trials in which fewer syllables were recalled	32.8	0.059	26.6	0.056
Difference between 1 and 3	−18.7		+14.0	

(32 subjects)

1. Subjects who recalled more syllables following 2-minute pause than immediately	28.1	0.079	53.1	0.088
2. Subjects who recalled the same number	28.1	0.079	21.9	0.073
3. Subjects who recalled fewer	43.8	0.088	25.0	0.077
Difference between 1 and 3	−15.7		+28.1	

* The formula for the $\sigma_{diff.}$ which involves the correlation of the two sets of data was employed where permissible. These σ's will be indicated throughout the tables by an asterisk (*). The correlations were obtained after pooling the determinations for each subject.

between the two methods of learning is also reflected in the calculations based upon percentage of trials and subjects showing reminiscence. Two sessions at the same stage of practice but involving different experimental programs are compared. These are called paired trials. The percentages

of paired trials show the frequency with which the recall after the rest pause is superior, equal or inferior to its control where the recall was immediately after the presentation. The data are pooled for the two lists learned with each condition by the thirty-two subjects for the comparison of subjects.

Reminiscence Following Partial Learning.—The effect of the two-minute rest pause with the two types of learning is shown at a later stage of practice in Table 2. The results

TABLE 2

RETENTION WITH AND WITHOUT 2-MINUTE REST PAUSE FOLLOWING LEARNING TO CRITERION OF 6 OUT OF 9 SYLLABLES WITH PAIRED-ASSOCIATES AND 7 OUT OF 12 SYLLABLES WITH SERIAL LEARNING

(*A*) *Recall Scores*

(32 subjects, 2 determinations for each condition)

	Paired-Associate		Serial	
	Mean	σ_M	Mean	σ_M
No. syllables correctly recalled immediately after criterion trial	6.12	0.141	7.02	0.172
No. syllables correctly recalled 2 minutes after criterion trial	5.44	0.157	7.53	0.181
Difference	−0.68		+.51	
$\sigma_{diff.}$*	0.164		0.218	
C.R.$_\sigma$	4.15		2.34	

(64 paired cases)

	Per cent	σ_P	Per cent	σ_P
1. Paired trials in which more syllables were recalled following 2-minute period than immediately	18.8	0.049	48.5	.062
2. Paired trials in which same number of syllables were recalled	29.7	0.057	26.5	.055
3. Paired trials in which fewer syllables were recalled	51.5	0.062	25.0	.054
Difference between 1 and 3	−32.7		+23.5	

(32 subjects)

1. Subjects who recalled more syllables following 2-minute pause than immediately	31.2	0.082	65.6	0.084
2. Subjects who recalled the same number	12.5	0.058	15.6	0.064
3. Subjects who recalled fewer	56.3	0.088	18.8	0.069
Difference between 1 and 3	−25.1		+46.8	

TABLE 2—*Continued*

(*B*) *Relearning to Criterion*
(32 subjects, 2 determinations for each condition)

	Mean	σ_M	Mean	σ_M
Trials required to relearn, no rest pause after criterion trial	0.48	0.112	0.56	0.083
Trials required to relearn, 2-minute rest pause after criterion trial	0.96	0.147	0.06	0.118
Difference	−.48		+.50	
$\sigma_{diff.}$*	0.171		0.096	
C.R.$_\sigma$	2.80		5.21	

(*C*) *Trials Required to Proceed to Mastery*
(32 subjects, 2 determinations for each condition)

	Mean	σ_M	Mean	σ_M
Trials required to proceed to mastery, no rest pause after criterion trial	6.51	0.551	9.05	0.671
Trials required to proceed to mastery, 2-minute rest pause after criterion trial	6.08	0.483	5.00	0.312
Difference	+.43		+4.05	
$\sigma_{diff.}$*	0.519		0.539	
C.R.$_\sigma$	0.83		7.52	

(64 paired cases)

	Per cent	σ_P	Per cent	σ_P
1. Paired trials requiring fewer trials to learn to mastery following 2-min. period than immediately	51.5	0.062	62.5	.060
2. Paired trials, same number required	9.4	0.036	23.5	.053
3. Paired trials, more required	39.1	0.061	14.0	.044
Difference between 1 and 3	+12.4		+48.5	

(32 subjects)

1. Subjects requiring fewer trials to learn to mastery following 2-min. pause than immediately	56.3	0.088	71.9	0.079
2. Subjects requiring same number	12.5	0.058	15.6	0.064
3. Subjects requiring more	31.2	0.082	12.5	0.058
Difference between 1 and 3	+25.1		+59.4	

for the serial learning group are reprinted from a previous article (**4**) for purposes of comparison. They were obtained with all experimental conditions the same. These series were twelve syllables in length. The rest pause was inserted after seven syllables were correctly anticipated. The criterion for the paired-associates was six out of nine syllables correct.

The comparability before the introduction of the rest pause of the groups tested with and without the rest pause is indicated by the following data:

Paired-Associates	Group I (Tested immed.)	Group II (Tested after two min.)
Trials to reach criterion of six syllables correct	9.47	9.64
No. syllables correct on criterion trial	6.34	6.26
Serial Learning		
Trials to reach criterion of seven syllables correct	6.39	6.18
No. syllables correct on criterion trial	7.48	7.52

With paired-associate material there is a loss in the number of syllables recalled when two minutes are introduced between the criterion trial and the subsequent recall (Table 2). Loss is also shown in relearning to the criterion performance following the rest pause. Since most subjects recall as many syllables or more on the trial following the criterion at this stage of learning, the distributions are badly skewed and the data are presented as of only incidental interest, and are not treated in great detail. A break-down of the data into paired trials and subjects, however, revealed the same results as the averages presented. The only evidence of reminiscence is the number of trials required to go from the criterion of six syllables correct to mastery. Even here, however, the improvement is not statistically significant, the critical ratio being only 0.83. In the case of serial learning, reminiscence is obtained with every index employed, and is statistically significant in the case of relearning to the criterion and in the number of trials required to proceed from the criterion performance to mastery.

These differences in the amount of reminiscence can be seen very clearly from Figs. 1 and 2 where the average numbers of repetitions required to learn varying numbers of syllables are presented for the paired-associate and serial learning. Figure 2, showing results on serial learning, is reproduced for comparison from a previous article (4) and the data were obtained under exactly comparable experimental conditions. The reminiscence in serial learning (Fig. 2) is in marked contrast to the complete absence of the phenomenon with paired-associates (Fig. 1).

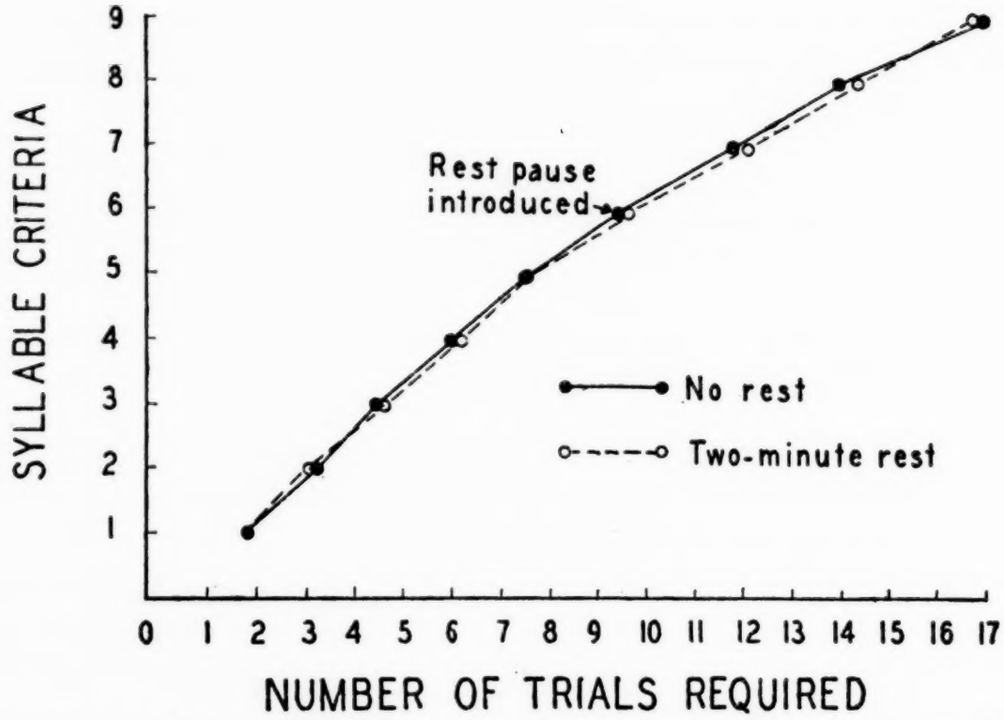


FIG. 1. *Paired-associate learning.* Composite curves showing average number of repetitions required to learn varying numbers of syllables with and without a 2-minute rest pause following the attainment of 6 syllables correct. Each point is based upon the mean of 64 learning scores.

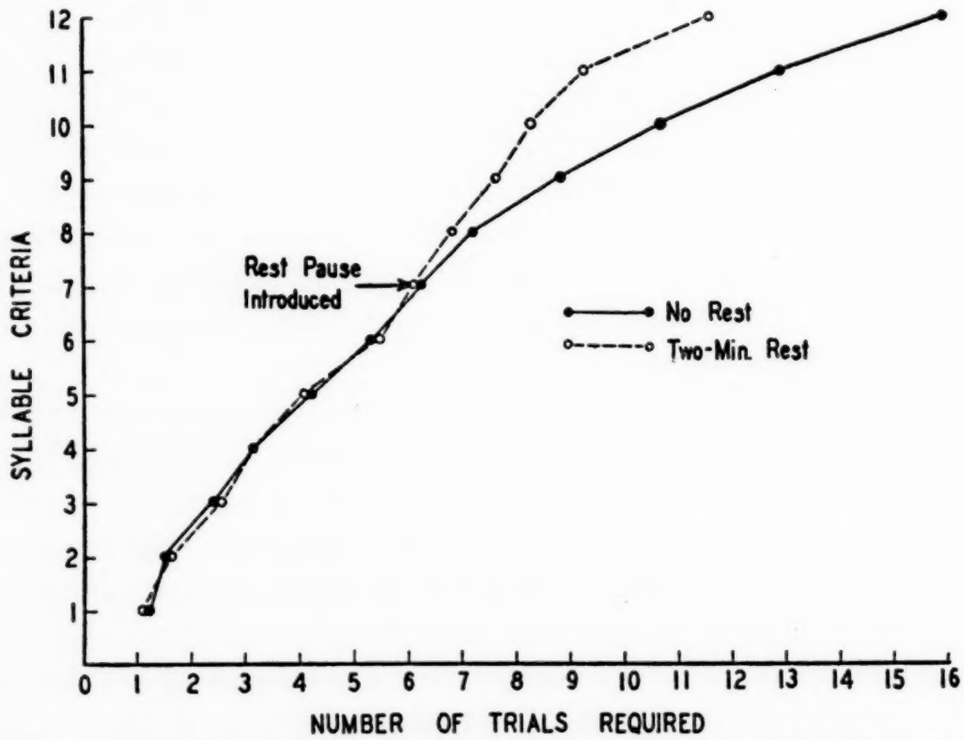


FIG. 2. *Serial learning.* Composite curves showing average number of repetitions of list required to learn varying numbers of syllables, with and without a 2-minute rest pause following the attainment of 7 syllables correct. Each point is based upon the mean of 64 learning scores. (Reprinted from Hovland (4).)

Reminiscence Following Mastery.—Results concerning the amount of reminiscence obtained with the two methods following the first perfect recitation trial are presented in Table 3.

TABLE 3

RETENTION WITH AND WITHOUT 2-MINUTE REST PAUSE FOLLOWING FIRST PERFECT RECITATION TRIAL WITH LEARNING BY SERIAL AND PAIRED-ASSOCIATE METHODS

(*A*) *Recall Scores*

(32 subjects, 2 determinations for each condition)

	Paired-Associate		Serial	
	Mean	σ_M	Mean	σ_M
No. syllables correctly recalled immediately after first perfect recitation	8.12	0.079	11.19	0.132
No. syllables correctly recalled 2 minutes after first perfect recitation	8.06	0.077	11.56	0.128
Difference	−0.06		+0.37	
$\sigma_{diff.}$*	0.099		0.161	
C.R.$_\sigma$	0.61		2.30	

(64 paired trials)

	Per cent	σ_P	Per cent	σ_P
1. Paired trials in which more syllables were recalled following 2-minute period than immediately	34.4	0.060	26.6	0.056
2. Paired trials in which same number of syllables were recalled	26.6	0.056	57.9	0.062
3. Paired trials in which fewer syllables were recalled	39.0	0.061	15.5	0.045
Difference between 1 and 3	−4.6		+11.1	

(32 subjects)

1. Subjects recalling more syllables following 2-minute pause than immediately	37.5	0.086	34.4	0.084
2. Subjects recalling same number	15.6	0.065	56.2	0.088
3. Subjects recalling fewer	46.9	0.088	9.4	0.051
Difference between 1 and 3	−9.4		+25.0	

(*B*) *Relearning to Criterion*

(32 subjects, 2 determinations for each condition)

	Mean	σ_M	Mean	σ_M
No. trials to relearn to mastery without rest pause following first perfect recitation	1.23	0.116	0.79	0.167
No. trials to relearn to mastery with 2-minute rest pause after first perfect recitation	2.06	0.177	0.40	0.145
Difference	−0.83		+0.39	
$\sigma_{diff.}$*	0.180		0.191	
C.R.$_\sigma$	4.61		2.04	

TABLE 3—*Continued*

	Per cent	σ_P	Per cent	σ_P
(64 paired trials)				
1. Paired trials requiring fewer trials to learn to mastery following 2-minute period than immediately	32.8	0.059	32.8	0.059
2. Paired trials, same number required	20.3	0.050	51.6	0.063
3. Paired trials, more required	46.9	0.062	15.6	0.046
Difference between 1 and 3	−14.1		+17.2	
(32 subjects)				
1. Subjects requiring fewer trials to learn to mastery following 2-minute period than immediately	34.4	0.084	37.5	0.086
2. Subjects requiring same number	18.7	0.069	50.0	0.088
3. Subjects requiring more	46.9	0.088	12.5	0.058
Difference between 1 and 3	−12.5		+25.0	

It will be recalled that because of trial-to-trial fluctuations, the average performance on the trial following mastery is almost invariably poorer than on the criterion trial (**14**). This decline was lessened by the introduction of a rest pause with serial learning. (Data of Table 3 on serial learning reproduced from (**4**).) In the case of paired-associates, however, recall scores, as well as relearning trials, indicate poorer performance following the introduction of a rest pause. Analyses based on percentages of paired cases and subjects likewise show forgetting of paired-associates with the passage of time, but reminiscence with serial learning. The percentages of trials and subjects showing improvement in relearning are almost identical for the two methods, but with paired-associates the percentages showing loss are much greater than with serial learning.

Discussion

A discrepancy between the present results and those of previous investigators is immediately apparent. Ballard (**1**), G. McGeoch (**10**) and others have previously obtained reminiscence with paired-associates. More recently McGeoch, McKinney and Peters (**12**) have confirmed these earlier studies. They found that the recall at the completion

of learning was 10.71 syllables, whereas after ten minutes the recall score was 11.48 syllables. This gives a ratio of delayed recall to original recall of 107.2 percent, cited by these authors as evidence for reminiscence. Superficially, then, these results are not in accord with the ones obtained in the present experiment, where, with paired-associate material, recall was poorer following the rest pause than immediately following learning. Some of the principal differences in experimental conditions between the studies of McGeoch, McKinney and Peters, and those of the present experiment should be mentioned:

(1) In the McGeoch, McKinney and Peters experiment paired Chinese-English words were used, whereas in the present study nonsense syllables were employed.

The evidence to date shows no clear-cut differences between different types of learning material in the conventional reminiscence experiment: "The question as to the material most susceptible to reminiscence awaits further experimentation" (G. McGeoch, **10**, p. 75). This problem has not been separately investigated for the short-time reminiscence studied by Ward and the writer.

(2) In the McGeoch, McKinney and Peters experiment, the subjects read over the lists four times at their own speed; in the present experiment, the pairs were automatically presented for two seconds each until the selected criterion of performance was reached.

(3) The subjects used by McGeoch, McKinney and Peters were allowed four minutes in which to write the English equivalent of the Chinese words presented by the method of right associates, whereas our subjects anticipated verbally the paired-associates as they were automatically presented.

While G. McGeoch has touched upon the problem of the effect of the method of presentation and recall, no adequate experimental data on the subject are available. Recently, in an unpublished experiment by Deese and the writer, it was found that when the subjects were allowed to read over the lists at their own speed for a given period of time and to write the material recalled, no reminiscence was obtained with

meaningful verbal material. But these results do not explain the discrepancy between the present results and those of McGeoch, McKinney and Peters, since they would tend to indicate that reminiscence is more apt to occur under conditions when syllables are presented repetitively at a constant rate than when an individual is given a period of time to learn them at his own rate.

(4) In the present experiment reminiscence was tested after two minutes; in their experiment ten minutes was used as the rest interval.

Ward has shown that ten minutes gives considerably less reminiscence than two minutes under his experimental conditions. This factor, however, would again work contrary to the discrepancy between the results, since they obtained reminiscence with the longer interval, whereas no reminiscence was obtained in the present study with the shorter interval.

(5) In the experiment of McGeoch, McKinney and Peters the orders of presentation were always the same, while in the present study the pairs were rotated on successive presentations.

This factor may be of considerable importance since, if the order of presentation is kept constant, long associational bonds may be formed which closely simulate those of serial learning. This point will be discussed below.

(6) Number cancellation was used as the interpolated activity during the rest pause by McGeoch, McKinney and Peters, whereas color-naming was employed in the present experiment.

The effects of different types of interpolated activity during the rest period upon reminiscence have not been separately investigated, but in the experiments of Ward (**14**), and of Deese and the writer, incidental results indicate that the presence or absence of reminiscence is not dependent upon the differences between various types of interpolated activity.

(7) For the principal results in the present study learning was carried on to a given criterion of performance while in the McGeoch, McKinney and Peters study the recall was measured after a given learning time.

The trial-to-trial fluctuations will result in a generally lower level of performance on the trial immediately after the criterion (see Ward, **14**). This would mean that with the same method of measurement, reminiscence would be less likely to occur when learning was to a criterion score. But the method of measuring reminiscence employed in the present experiment takes account of this by using as a control the performance on a trial immediately after the criterion. The effect of this difference in procedure is discussed in the succeeding paragraphs.

(8) In the McGeoch, McKinney and Peters study, the recall after a rest period is compared with the recall at the conclusion of learning; in the present experiment and in others reported by Ward and the writer, the amount recalled when a period of time elapses between learning and recall is compared with the amount recalled on a control session in which the recall is tested on the trial immediately following the learning.

This difference in procedure seems to the writer the crucial factor in explaining the discrepancy between the results of McGeoch, McKinney and Peters and those here reported. If reminiscence is to mean 'the improvement in the recall of incompletely learned material after an interval of time without intervenient formal relearning or review' (**10**) it would seem most logical to measure it by comparing the performance after the rest pause with a control trial also following the learning but with a zero time interval between learning and recall. The results will not then be complicated by additional practice. For an elaboration of this point, the reader is referred to the article by Ward (**14**). He points out that the last learning trial 'is not a real measure of the degree of learning.' It is 'only a measure of the degree of learning at the end of the previous trial and after a time interval equal to the length of rest (if any) between trials.' . . . "Any changes in score from one trial to the next include not only changes due to time factors, but also changes *due to the practice* [1] received during the former trial" (p. 20). It appears, therefore, that Mc-

[1] Italics the present writer's.

Geoch, McKinney and Peters, as well as G. McGeoch and Ballard, have not demonstrated reminiscence with paired-associates according to the usual definition of the term. Their results might be explained as being due to the effect of recitation upon recall, a factor shown to be of considerable importance not only in learning (Witasek (**15**) and others) but also in *recall* (Myers, **13**). The delayed recall scores in the former studies of reminiscence always represented recall following an additional practice period, even if only a single trial in length. This possible explanation of typical reminiscence results was mentioned by G. McGeoch, but discarded because of lack of evidence. The implications of this factor will be discussed in later articles in the present series by the writer.

The main theoretical problem remains, *i.e.* that of explaining the absence of reminiscence with paired-associate materials with experimental conditions identical to those under which pronounced reminiscence had been obtained by the writer and by Ward with serial learning. The principal theories of reminiscence previously reviewed (Hovland (**4**)) shed little light on these results. Some of them are clearly inapplicable. *Rehearsal* during the rest interval, for example, could not have had a differential effect, since in both cases rehearsal was prevented by color-naming during the interpolated rest pause. *Motivation* as an explanation lacks any experimental confirmation. *Perseveration* could be used to explain the results only if it could be shown that perseveration effects are greater with serial than with paired-associate materials. This has not been demonstrated.

Reminiscence phenomena have also been explained on the basis of *fatigue*. Fatigue in the physiological sense would not appear to be a good explanation for the results under discussion, since reminiscence is obtained with serial learning and is not obtained with paired-associates even after the number of trials has been equated. Fatigue would presumably vary with the amount of time spent in the learning.

A closely similar type of explanation is one based upon decrement on the analogy of *refractory phase* phenomena. Decrement might be expected to occur when the stimulus-

response sequences occurred in rapid succession. Recovery from the decremental effect would then be expected with the passage of time. To use this type of explanation for the present results we would have to assume that decremental effects due to refractory phase were less prominent in paired-associate than in serial lists. Two possible bases for such an hypothesis could be given: (1) It might be that the decrement is dependent not upon the rate of the stimuli, which were at two-second intervals in both methods of learning, but upon the rate of responses by the subject, which are made once every two seconds with serial learning, but only once every four seconds with the paired-associates. It will be recalled from the section on experimental procedure that two seconds were allowed for each member of the pairs to correspond with the two seconds given for each syllable in the serial method. The subject responds to each alternate syllable in this method so that his responses occur at four-second intervals, while in serial learning the subject responds to every syllable and this means that his responses occur at two-second intervals. This explanation is strengthened by the fact that the phenomena obtained with paired-associate materials so closely simulate those obtained with the four-second time interval in serial learning. Clear-cut evidence as to this factor is difficult to obtain, however, because with shorter time intervals, the amount of time devoted to the individual components would aggregate less than that in the serial method and hence prevent comparability. It is interesting to note in passing, however, that the four-second speed required far fewer trials for learning than the two-second rate in serial learning, but that even with the present timing of two seconds per element of the paired-associate list, it requires as long a time to learn nine paired-associates as eleven or twelve syllables in serial learning. (2) The other alternative explanation would be that the continuous rotation of the order of the pairs and the alternate exposure of the syllables in different windows of the apparatus is less monotonous to the subject, so that decremental effects are less in paired-associate lists than in serial learning where the syllables are presented in the same order in

the same window and at the same rate. This type of explanation is based upon attitudinal changes and is at a much higher level than that at which refractory phase phenomena are obtained.

Theories based upon 'inhibition' or 'interference effects' are somewhat related. As suggested above, these theories would see in the difference in the structure of the learning with the two methods, particularly in the difference in the frequency and character of remote associative connections, the explanation for reminiscence. The present experiment, taken in conjunction with the results of Kjerstad, cited above, suggests considerable merit in this type of explanation. But greater predictive efficiency of this theory must still be demonstrated. A possible experiment to test the differential predictions of the refractory phase and interference theories would be to study reminiscence with paired-associates in which the order of presentation of the successive pairs was not varied. This procedure would permit formation of long associative bonds between the successive pairs. If the interference theory is correct, it would be predicted that the proposed experiment would show reminiscence effects intermediate between the present results and those for serial learning.

Summary and Conclusions

1. To study possible reminiscence effects in paired-associates, retention following a two-minute rest period between learning and recall was compared with retention on control trials immediately following learning for the same subjects. The experimental conditions were maintained the same as those employed in earlier studies by the writer on serial learning (**3, 4**) to afford comparison of reminiscence with these two learning methods. The lists learned consisted of nine pairs of nonsense syllables. The successive pairs were presented in constantly changed order to minimize the formation of long associational bonds. Recitation during the rest period was prevented by having the subjects name colors presented automatically.

2. Instead of the reminiscence effect obtained in serial

learning, forgetting was found with paired-associates during the two-minute rest pause following a single presentation of the list. Fewer paired-associates were correctly recalled on the first recitation trial when it occurred after two minutes than when it occurred immediately after learning; with serial learning, on the other hand, it had been found that more syllables were recalled on the trials when the rest pause was introduced.

3. Following learning to a criterion of six out of nine associates correct, forgetting was likewise obtained during the two-minute rest pause with paired-associate materials. The only suggestion of reminiscence obtained was a reduction in the number of trials required to go from the criterion to mastery following the rest pause. Pronounced reminiscence had been obtained with all of these measures with serial learning.

4. The same difference between paired-associate and serial learning was shown on the recall and relearning with and without a rest pause following the first perfect recitation. Paired-associate material showed decreased recall and more relearning trials ('forgetting'), while serial learning has been shown to give reminiscence under these circumstances.

5. The discrepancy between the present results and those of other investigators who obtained reminiscence with paired-associate materials is discussed in the light of differences in procedure. The difference in results is attributed primarily to the difference in the method of measurement employed. Customarily, reminiscence has been claimed when a second recall after a lapse of time is better than the first recall immediately after learning. This may really not be reminiscence as customarily defined but only improvement as a result of an additional recitation trial. In the present series of experiments and in Ward's study, the comparisons have been between the recall following a rest pause and the recall on a control series in which no rest pause was introduced. This rules out the possibility that any increased recall is due to an additional practice trial.

6. Theories of reminiscence customarily advanced are discussed and criticized in relation to the present experiment. It

appears most likely that differences in the type of decremental effects in the two methods of learning and/or the nature of the interferences involved are responsible for the results. An experiment to test the differential predictions of these two types of explanation is suggested.

(Manuscript received November 17, 1938)

References

1. Ballard, P. B., Obliviscence and reminiscence, *Brit. J. Psychol. Monog. Supp.*, 1913, 1 (No. 2), pp. vi + 82.
2. Ebbinghaus, H., *Über das Gedächtnis*, Leipzig: Duncker and Humblot, 1885, pp. ix + 169.
3. Hovland, C. I., Experimental studies in rote-learning theory. I. Reminiscence following learning by massed and by distributed practice, *J. Exper. Psych.*, 1938, **22**, 201–224.
4. Hovland, C. I., Experimental studies in rote-learning theory. II. Reminiscence with varying speeds of syllable presentation, *J. Exper. Psych.*, 1938, **22**, 338–353.
5. Hovland, C. I., Experimental studies in rote-learning theory. III. Distribution of practice with varying speeds of syllable presentation, *J. Exper. Psych.*, 1938, **23**, 172–190.
6. Hovland, C. I., Experimental studies in rote-learning theory. V. Comparison of distribution of practice efficacy with serial and paired-associate learning methods (in prep.).
7. Hull, C. L., The conflicting psychologies of learning—a way out, *Psychol. Rev.*, 1935, **42**, 491–516.
8. Hull, C. L., R. T. Ross, D. T. Perkins, and C. I. Hovland, *A Mathematico-Deductive Theory of Rote-Learning—a Study in Methodology* (in prep.).
9. Kjerstad, C. L., The form of the learning curves for memory, *Psychol. Monog.*, 1919, **26** (No. 116), pp. v + 89.
10. McGeoch, G. O., The conditions of reminiscence, *Amer. J. Psychol.*, 1935, **47**, 65–89.
11. McGeoch, J., The direction and extent of intra-serial associations at recall, *Amer. J. Psychol.*, 1936, **48**, 221–245.
12. McGeoch, J. A., F. McKinney, and H. N. Peters, Studies in retroactive inhibition. IX. Retroactive inhibition, reproductive inhibition and reminiscence, *J. Exper. Psych.*, 1937, **20**, 131–143.
13. Myers, G. C., Recall in relation to retention, *J. Educ. Psychol.*, 1914, **5**, 119–130.
14. Ward, L. B., Reminiscence and rote learning, *Psychol. Monog.*, 1937, **49** (No. 220), pp. v + 64.
15. Witasek, S., Über Lesen und Rezitieren in ihren Beziehungen zum Gedächtnis, *Zeitschrift f. Psychologie*, 1907, **44**, 161–185; 246–282.

FINGER OSCILLATIONS AS INDICES OF EMOTION I: PRELIMINARY VALIDATION *

BY F. K. BERRIEN

Colgate University

A host of studies by Barnacle, Ebaugh and Lemere (1), Burtt (2), Cason and Cason (3), Duffy (5), Kellogg (9), Krause (10), Laird (11), Luria (12), Morgan (13), Olson and Jones (14), Runkel (15), Skaggs (16), Gardner (6), Huston, Shakow and Erickson (8) have demonstrated that emotional experiences involve, or are accompanied by disturbances in the voluntary musculature. The purpose of the two experiments which are to be reported [1] was to explore the possibilities of using the variations or disturbances in a simple voluntary movement as an indicator of emotion. It was hoped that a technique might be developed which could be used to improve the existing methods of detecting deception by supplementing records of blood pressure, pulse, breathing and galvanic skin responses.

A review of the literature has revealed a poverty of information concerning what Dodge and Bott (4) call the 'movement pattern' as distinguished from the 'muscular' and 'neural' patterns of action. By movement pattern, Dodge and Bott mean that pattern of changes in the position and rate of movement of a given member which is brought about through muscular and neural events. Aside from Hollister's study (7) in the field of speech, there has been apparently no comprehensive attempt to ferret out the variables which are inherent in the movement pattern itself.

Finger movements were selected as the particular responses for study because they satisfy certain elementary requirements. (1) Such movements are relatively easy to

* The writer takes this opportunity to express his gratitude to Dr. H. E. Burtt of Ohio State University under whose direction the present investigation was conducted.

[1] The second experiment will appear at a later date.

execute. (2) They can be maintained without serious fatigue for periods of 30 minutes or more. (3) We believed, and later confirmed the belief that they could be reduced to an automatism within a relatively short time, thus leaving the subject relatively free for mental activity. (4) Though easy to execute, they involve minute adjustments which may be recorded and which may be of significance for this study. Another factor contributing to the decision to use finger movements was the existence of a number of recent studies among those indicated above which deal with digital responses as indicators of emotion.

Experimental Task and Apparatus

The essential task required of a subject was to flex his index and middle fingers in response to metronome signals. The subject sat in a chair provided with a broad right arm rest such as is ordinarily found in college lecture rooms. Attached to a short extension of this arm rest was a right angle response lever (Fig. 1, *A*) pivoted at the

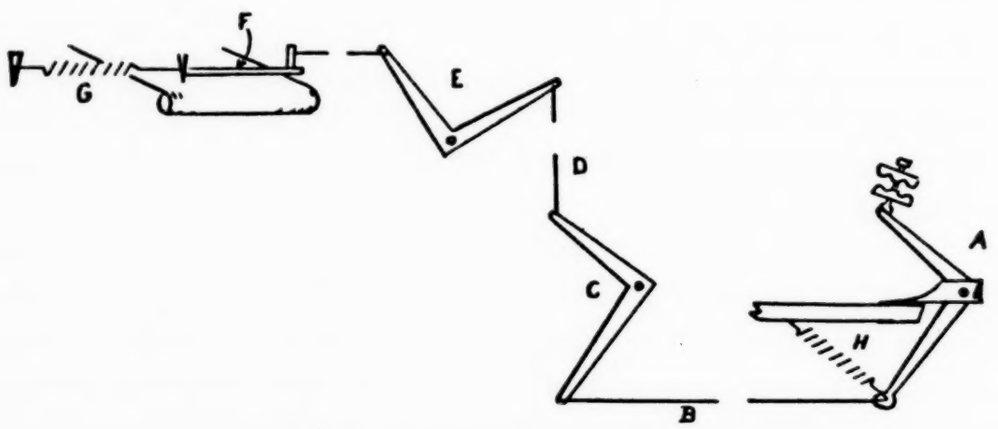


Fig. 1. Schematic diagram of digitalgraph. See text.

point of the angle so that it might rotate in the vertical plane. On one leg of this lever was attached a bakelite clamp shaped to fit the fingers. Preliminary experimentation revealed the necessity of connecting this clamp to the response lever itself by a ball and socket joint in order to prevent some slight discomfort which developed when the clamp was fastened rigidly to the lever. Since the movements were in a vertical direction the flexibility of the joint did not interfere with the accuracy with which the movements were transmitted. The pivot mounting was arranged so that the palmar surface of the fingers had to be flexed to a point about ¾ of an inch below the level of the rest of the hand before the limit of movement would be reached.

Attached to the other end of the right angle lever was a fish line (*B*) 1/32 of an inch in diameter of 18 strand braided silk coated with shellac. This line ran directly under the arm of the chair to another pivoted right angle lever (*C*) attached to the rear of the

chair arm. The fishline (*D*) then followed a vertical direction up to a third right angle lever (*E*) which served to transform the direction of movement back to the horizontal. The line running from the second leg of the third lever was connected with a pen bar (*F*) that moved back and forth across a 3 inch polygraph tape. In order to cause the pen bar to return to its base line when the fingers were being extended a spiral spring (*G*) of piano wire approximately three and a half inches long was connected by a silk thread to the opposite end of the pen bar. A little experimentation with the apparatus showed the necessity of including a second spring to increase the 'back pressure' of the system on the fingers. This spring (*H*) was connected to the second leg of the first right angle lever at the same point as the fish line. The coiled spring was anchored to the underside of the chair arm. The minimum pressure required to move the response lever from its resting position was 40 grams. The maximum pressure, that is, the pressure required to hold the lever at the opposite extreme, was 87 grams. The distance over which the fingers were permitted to move was 5.6 cm. Seldom if ever were oscillations made over this entire distance.

The subject was paced by a metronome mounted behind his chair. The metronome was equipped with mercury cups that served as electrical contacts for completing the circuit of an electromagnet. This magnet operated a pen rod mounted in such a way that an offset appeared in an otherwise straight line each time the metronome sounded. A similar magnet and pen were operated by closing a knife switch controlled by the experimenter to indicate when a stimulus was given and when the subject responded verbally. For convenience the apparatus has been called a digitalgraph.

Subjects and Procedure

Since it was the purpose of this experiment to discover those variables which changed consistently with the onset of an emotion it was thought best to use as subjects psychopathic patients whose histories were relatively well known. From these histories information could be derived concerning embarrassing and intimate incidents. It seemed reasonable to assume that if patients were questioned about such incidents they would more likely experience an emotion than when asked questions of a general nature such as, "Did you have breakfast this morning?"

The 16 subjects for this experiment were all in-patients of the State Hospital at Ypsilanti, Michigan, when the records were made. They varied in age from 19 to 65. The cases were restricted to males because of the nature of the questioning that was planned.

The instructions which the patients received were as follows:

> I am going to ask you to sit in this chair and keep time with this ticker or metronome by tapping with your fingers. This won't hurt you at all. Each time you hear a tick behind you just press down on this lever. It is not necessary that you press as far as you can go. Just a little movement will be sufficient, like this. (Experimenter moved subject's fingers in time with metronome signals over a distance of about 2 cm.) I am going to let you practice for a minute or two and then I will ask you some questions. No matter what I say, just keep right on tapping and answer my questions. All of the questions can be answered by saying either yes or no. If you want to add something to your answers just wait until after we stop the machine and then you can tell me about it. Any questions? Remember, this isn't going to hurt you one bit.

The questions were framed after a study of the case history of each patient. The number of questions asked each patient varied from eight to nineteen with an average

of twelve. In every instance two control or 'shock absorber' questions were given after the fingers had attained an even, regular rhythm. Two, three, or four questions were given concerning one topic that had played a prominent part in the history, such as extra-marital relations, vocational maladjustments, and feelings of guilt or inferiority. One or two additional control questions were then given with the idea of permitting the patient to return to normal. Another set of questions concerning a different aspect of the patient's case followed. This alternation of control and crucial questions continued throughout the record.

The rate of tapping was approximately 60 flexions per minute; a rate that some preliminary experiments with normal college students had indicated as the most advantageous.

Method of Analyzing Records

Seven characteristics of the records were measured and from these, two additional indices were derived. In order to expedite the analysis, the records were magnified approximately 9 times by projecting them upon a screen by means of a baloptican. The screen consisted of cross sectioned paper. Each situation or question was divided into two portions; the normal period consisting of ten cycles or taps just preceding a question, and the response period extending from the moment a question was begun to a point ten beats after the answer was given.

Length

The first characteristic of the curves studied was the average length of each cycle. This was derived by measuring the entire length of the response line for a period by means of a map measure and dividing by the number of cycles in the period. The values were expressed in inches.

T Units

The second characteristic examined was the average number of time units elapsing between the moment the upstroke was almost completed and the moment the downstroke was definitely under way. The time units (hereafter referred to as T units) were determined in the following manner: The maximum point reached by a cycle was determined on the cross sectioned screen. By dropping down one unit distance from the maximum it is then possible to count the number of units laterally between the upstroke and the downstroke of any given cycle. The procedure is illustrated in Fig. 2 which shows in cycle A four such time units and in B only two time units.

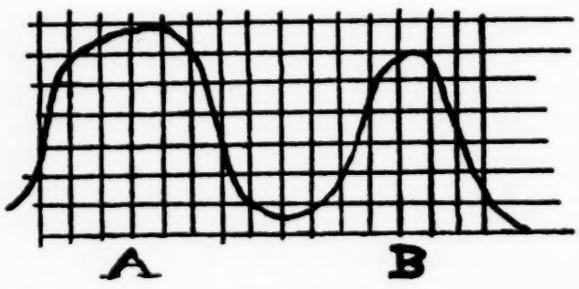


Fig. 2

It is not possible to say exactly what the absolute value of the T unit is, but each unit distance of the magnified records represents approximately .1 seconds.

Irregularities

The irregularities of the up or down strokes were divided into three classes illustrated in Fig. 3.

FIG. 3

Class I irregularities were defined as any slight deviation in the usual form of the curve. Class II irregularities were those which were characterized by a well defined plateau. Class III included alterations in the direction of movement. It is evident that such a classification reveals the 'seriousness' of the irregularity. Wherever two classes of irregularity occurred on the same stroke of the same cycle only the more serious was counted. Irregularities for the upstroke were counted separately from those for the downstroke. Thus the frequency of irregularities on each half of the cycle was kept separate.

An irregularity score for each half cycle was derived in the following manner: The frequency of each class of irregularity in a given period was determined. Class I frequencies were given a factor of 1, class II frequencies were given a factor of 2 and class III frequencies, a factor of 3. The frequencies were multiplied by these factors. The products were added and divided by the number of cycles in the period. Later the irregularity score for the upstroke was averaged with that for the downstroke to give an average irregularity score for the cycles.

F/E Ratio

The flexion-extension ratio for each period was found in the manner described below. The number of time units elapsing between the beginning of the downstroke and the minimum point on the curve was counted and became the numerator of the fraction. This value was divided by the number of horizontal units between what was obviously the beginning of the next upstroke and its maximum point.

Figure 4 illustrates the procedure. In the sample shown, eight time units elapse

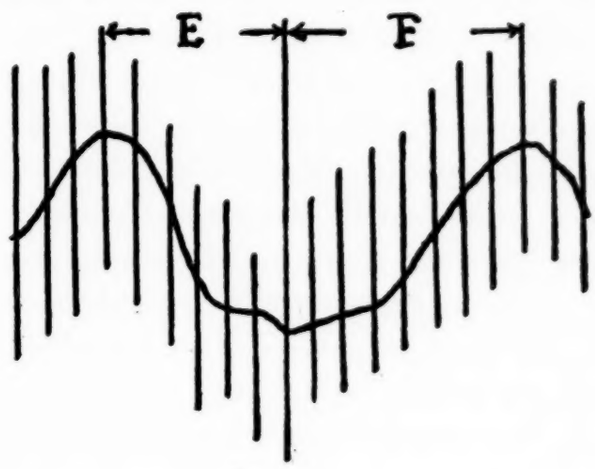


FIG. 4

between the beginning and the completion of a flexion movement. Only six such units are required for the extension movement. The F/E ratio is therefore 1.33.

The mean ratio for the normal period was based on only the 5 beats just preceding the response period. The mean ratio for the response period was based upon the ratios of all cycles between the beginning of the response period and a point 5 beats after the answer was given. Because a tabulation of the ratios revealed a skewed distribution the geometric mean was used as the measure of central tendency in both periods.

Variability

The average variation in the point at which the fingers changed their direction of movement was likewise examined. Let us consider variations in the point at which the fingers started the downstroke. This begins at the maximum point of the cycle. It is necessary merely to count the number of units in the vertical direction on the cross sectioned screen between each maximum point and its neighbor in order to have some measure of variability. The total number of vertical units in each period separating each maximum point from the next following without regard to direction up or down was determined and divided by the number of cycles. This value was known as crest variability. The corresponding value was determined for the troughs of the cycles and was called trough variability. A total variability score was derived by adding the average scores for crest and trough. Figure 5 illustrates the procedure.

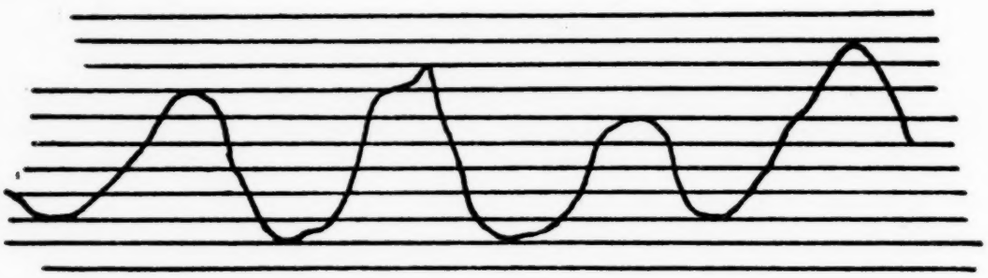


FIG. 5. Trough, right to left, 1, 0, 1
Crest, right to left, 3, 2, 1

RESULTS

Inspection of these data for the 120 crucial questions and the 70 controls revealed in the first place that no single characteristic changed consistently on the crucial questions and did not change on the controls. That is to say, when we compared the mean of a given variable (such as F/E ratio) in the normal period, with the mean for the same variable in the response period, there appeared to be no consistent tendency for such a variable either to increase or decrease in response to crucial questions. The variables, of course, did change from period to period, but the direction and magnitude of such changes did not appear to be related to the type of question asked.

Further inspection of the data however showed that three of the original nine variables were probably more indicative of emotion than the others. These three were average length per cycle, variability of the trough and T units.

The problem then became one of combining these three measures into a single index which would be diagnostic of emotion.

The usual methods of combining scores of unlike variables by the use of standard scores could not be employed in this instance because of one serious difficulty. In the case of length of cycle the method of measurement involved getting only the sum of the cycles rather than measuring each individual cycle. From this total the average was obtained. Consequently, no measure of variability around the mean was possible for this variable. Hence standard scores were impossible to obtain.

The alternative that was chosen involves the use of index numbers. This procedure makes the assumption that the control period of each question is the standard with which subsequent changes in the response to that question are to be compared. The procedure essentially equates the values of each control period, in that each is assigned the value of 100. The response period average of a given variable is divided by the control period average and the quotient expressed as a percent. This gives essentially an indication of the extent to which the response period value of a given variable exceeds, or is exceeded by, the normal period value. The stated degree of change is relative to the absolute value of the normal period.

The indices for average length, T units, and variability of the troughs were derived for all 190 questions included in this experiment. From these data three frequency tables were compiled each showing the distribution of one of the indices in response to control and crucial questions. Each of these tables represents 190 responses classified in respect to two characteristics. One of these characteristics is measured by a continuous variable (average length index, for example) and

the other characteristic is a dichotomous variable (crucial or control questions).

At this point a number of possibilities present themselves as ways of combining these indices, based on the three measured characteristics. In the first place, the indices for length, T units, and variability might be added together for a given question. This would have the effect of weighting the variables equally. Even a casual inspection of the data shows, however, that these variables are not equally effective in revealing emotional responses.

Secondly, weights might be assigned to the variables arbitrarily on the basis of inspection of the data. Numerous such attempts at guessing weights might eventually yield a combination which would be reasonably satisfactory, but it would be impossible to tell when further guessing would be useless.

Thirdly, some general notion of the optimal weights could be obtained by computing the zero order correlations between each of the variables and the response we are attempting to detect, namely emotion. Large differences in correlation would suggest at least some difference in weight. The response in question is not reducible to a quantitative score in our data as we merely have control and crucial questions which are assumed to give in one case non-emotional and in the other, emotional reactions. Introspective gradations of emotional responses did not appear feasible with the untrained subjects. For this reason the regular Pearson *r* could not be employed to obtain zero order correlations.

The biserial *r* technic appeared to be applicable in this instance. Such *r*'s should give some notion of the relationship existing between these variables and the presence or absence of emotion. If the intercorrelations likewise are computed it would be logical to give greater weight to a measure that had a high correlation with the emotional factor and a low correlation with the other measures.

Finally, these weights will more closely approach the optimum if the partial correlation procedure is applied to the zero order coefficients. Such a procedure culminates in a

regression equation whose coefficients are the weights to be assigned. This procedure enables us to assign weights to the variables with a minimum of guessing and a maximum possibility that the weights approach an optimal relationship.

It must be remembered, however, that the dependent variable involves only two categories and that our regression equation cannot predict degree of emotion as might be possible if the original data involved a continuous gradation of the emotional factor. Our interest is in deriving a composite index such that very atypical indices will be indicative of emotion.

While an index derived in the manner just described may not be the best possible, nevertheless, in the light of partial correlation theory it is highly probable that it will be superior to any of the other techniques just recounted.

The combination of weights derived by partial correlation and taken to the nearest whole number is represented in the equation

$$X = 32L + V - 5T,$$

where X equals the combined index, L the index value for length, V the index for variability of trough, and T equals the index for T units.

The distribution of these X values computed according to the above equation is given in Table 1. It will be observed that few of the crucial responses exceed either limit of the controls as presented in this table. However, this fact does not mean that there were few crucial responses which received scores that deviated beyond the limits of the *individual's* controls. On the contrary, if the composite indices obtained by a given individual are separated into controls and crucial responses we find many crucial responses which deviate more than four average deviation units from the mean index of the controls.

The criterion of four A.D. units spread from the control mean was accepted as indicative of a significant departure from that mean. A departure of such great magnitude appeared to establish beyond doubt that the observed value is a member of a distribution other than the control distribution.

TABLE 1

DISTRIBUTION OF COMPOSITE INDICES WHERE $X = 32L + V - 5T$

Composite Indices	Frequency Control Responses	Frequency Crucial Responses	Frequency Atypical Responses
5800–5899		1	1
4800–4899		1	1
4000–4099	2		
3900–3999		1	1
3800–3899		1	1
3700–3799		6	3
3600–3699	1	4	2
3500–3599		3	
3400–3499	2	2	1
3300–3399	5	9	
3200–3299	2	7	2
3100–3199	8	12	
3000–3099	8	16	3
2900–2999	9	8	
2800–2899	7	11	
2700–2799	8	11	
2600–2699	4	6	
2500–2599	5	7	
2400–2499	6	6	
2300–2399	1	3	1
2200–2299		1	
2000–2099	1	2	
1700–1799		1	1
1400–1499		1	1
Total	70	120	18

Indices which were 4 A.D. units from the control mean hereafter will be referred to as *atypical indices*.

In the last column of Table 1 the responses which meet this criterion are listed in the appropriate class. It will be noted that 15 percent of the 120 crucial responses are indicated in this column.

Several other methods of treating the data were attempted without significant results.

DISCUSSION OF RESULTS

The fact that only 15 percent of the responses to crucial questions are classified as emotional by the criterion referred to above is not to be dismissed readily. The entire experiment was undertaken with the belief that many of the questions classified as crucial would not excite emotional responses. It must be remembered that all of the patients had undergone

previous medical and psychiatric examinations during which the intimate details of their lives had been discussed freely. In many cases this discussion had taken place before a mixed group including the medical, psychological and social service departments of the hospital. The emotional potency of the questions was considerably reduced by having had this previous experience and also by the fact that the finger records were all taken with only the experimenter present. This conclusion is borne out by the testimony of some of the more cooperative patients who admitted that many of the questions affected them to only a slight degree because they were aware that the experimenter knew the details of their stories.

The fact that none of their responses to control questions were accompanied by indices that deviated beyond 3 A.D. units of the control mean for a given subject, lends additional evidence to the validity of the procedure. The data, therefore, appear to indicate that the presence of an atypical index may be taken as diagnostic of emotion but the absence of such an index does not necessarily imply the absence of emotion.

Inspection of Table 1 shows that many, but not all of the responses to crucial questions which are atypical have scores that are *greater* than the controls. Three of the responses were significantly *less* than the controls. The existence of these three cases lends support to the belief that there are at least two types of responses which are atypical. The first and more frequent is characterized by an increase in the length of cycle, an increase in variability and a decrease in *T* units. This type might be called excitatory since all of the indices reveal an increase in activity.[2] These responses yield high composite indices.

The second type of response is almost the reverse of the first. There is a decrease in length, a marked increase in T units, and sometimes an increase in variability. Such records show a decrease in activity. They can be described as

[2] A decrease in T units means the fingers are at rest a shorter time between any two cycles.

the inhibitory type and usually receive low composite indices.[3]

After the above hypothesis had been formulated on the basis of the composite index data alone, it was recalled than an entire record made by a dementia praecox patient had been discarded because the finger responses had practically ceased on almost every crucial question. It was obviously meaningless to attempt an analysis of such a record according to the usual methods. In the light of what has been outlined it became clear that this patient was exhibiting an extreme form of the inhibitory type of response.

Returning to a consideration of the results, there appears to be another reasonable explanation for the rather small number of responses to crucial questions which are classified as atypical. There is, in the first place, a respectable body of fact which generally discredits psychological types. The general conclusion reached by the critics of 'type psychology' is to the effect that the so-called types are the extremes of the normal distribution. The majority of people, as well as the majority of an individual's reactions, fall within the interquartile range on almost any scale that is set up or in any performance that is measured.

The logical corollary to the assumption of two basic kinds of responses in the light of the above paragraph seems to be that the inhibitory and excitatory tendencies are extremes, and more often exist at approximately the same time, thus cancelling one another. Since it is also true that the variables involved in determining the composite index have very low intercorrelations, it may well be that on some questions one variable reflects the inhibitory tendency and another variable reflects the excitatory tendency. In such instances the composite index would of course not reveal either tendency clearly.

The hypothesis is strengthened by a study of the data. Of the three responses which were of an inhibitory type two came from the record of a drug addict who also gave one atypical excitatory response. The other was given by an

[3] Our use of the terms excitatory and inhibitory in no way implies any assumption concerning the neurological factors involved. The terms merely describe the production of the subjects.

hysteric who also gave two other excitatory responses. For this reason it is not possible at this time to conclude that the two kinds of responses are confined to personality types. The evidence on this point, however, is somewhat meager. The question can perhaps be answered by further experimentation. On the other hand, there can be no doubt that the two types of responses exist and at times are clearly revealed by the composite index. Moreover, since both kinds of responses appear in the same individual on different questions it is reasonable to assume that on other questions the two tendencies may exist together, thus being obscured in the composite index.

Summary and Conclusions

The present experiment is the first of two which have sought to validate a method of analyzing digitalgraphic records for purposes of detecting emotion. Rhythmical finger movements of psychotic patients made while being questioned about embarrassing details of their histories were reduced to nine indices based on objective characteristics. Three of these indices were combined by means of partial correlation technique, thus giving a composite index. When some crucial questions were asked the index deviated markedly from the mean of the indices obtained in response to control or nonemotional questions. While only 15 percent of the crucial questions were accompanied by such atypical indices, a number of conditions were cited which in part, at least, account for the low efficiency of the procedure.

There was evidence of two extreme types of reaction in the presence of emotion. One of these was described as inhibitory, and the other as excitatory in nature. The principal conclusion which follows from this experiment is that an atypical composite index is probably a valid indication of the existence of an emotional response, but the absence of an atypical index does not necessarily imply the absence of an emotional response.

(Manuscript received November 12, 1938)

References

1. Barnacle, C. H., Ebaugh, F. G., and Lemere, F., Association-motor investigation of the psychoneuroses, *Amer. J. Psychiat.*, 1935, **91**, 925–937.
2. Burtt, H. E., Motor concomitants of association reaction, *J. Exper. Psychol.*, 1936, **19**, 51–63.
3. Cason, H. and Cason, E. B., Affectivity in relation to breathing and gross bodily movement, *J. Gen. Psychol.*, 1933, **9**, 130–156.
4. Dodge, R. and Bott, E. A., Antagonistic muscle action in voluntary flexion and extension, *Psychol. Rev.*, 1927, **34**, 241–272.
5. Duffy, E., Tension and emotional factors in reaction, *Genet. Psychol. Monog.*, 1930, **7**, No. 1.
6. Gardner, J. W., An experimental study of the Luria technique for detecting mental conflict, *J. Exper. Psychol.*, 1937, **20**, 495–506.
7. Hollister, R. D. T., Relation between hand and voice impulse movements, *Speech Monog.*, 1937, **4**, 75–100.
8. Huston, P. E., Shakow, D., and Erickson, M. H., A study of hypnotically induced complexes by means of the Luria technique, *J. Gen. Psychol.*, 1934, **11**, 65–97.
9. Kellogg, W. N., The effect of emotional excitement upon muscular steadiness, *J. Exper. Psychol.*, 1932, **15**, 142–166.
10. Krause, L. S., Relation of voluntary motor pressure disorganization (Luria) to two other alleged complex indicators, *J. Exper. Psychol.*, 1937, **21**, 653–661.
11. Laird, D. A., Changes in motor control and individual variations under influence of 'razzing,' *J. Exper. Psychol.*, 1923, **6**, 236–246.
12. Luria, A. R., *The nature of human conflicts*, Liveright, 1932.
13. Morgan, J. J. B., The overcoming of distraction and other resistances, *Arch. Psychol.*, 1916, **5**, 1–84.
14. Olson, D. M. and Jones, V., An objective measure of emotionally toned attitudes, *J. Genet. Psychol.*, 1931, **39**, 174–196.
15. Runkel, J. E., Luria's motor and word association in the study of deception, *J. Gen. Psychol.*, 1936, **15**, 23–37.
16. Skaggs, E. B., Changes in pulse, breathing and steadiness under conditions of startledness and excited expectancy, *J. Comp. Psychol.*, 1926, **6**, 303–318.

STUDIES IN HIGH SPEED CONTINUOUS WORK: I. PERIODICITY *

BY B. R. PHILIP

Queen's University

Most studies of high speed continuous work not almost purely mental in nature do not go beyond 20 mins. in the one session. The work curves obtained during this time interval show a pronounced decrement almost from the beginning of the work period. The curves have been fitted by Ioteyko (1) to the third degree parabola, $y = a - bx + cx^2 - dx^3$, and by Neifeld and Poffenberger (2) to the parabola, $y = a + bx - cx^2 + dx^3$. Philpott has fitted curves to the hyperbola, $y = k \log \frac{t + a}{t - a} + C$ (3). In none of these studies of relatively short periods of work is there any evidence of periodicity, though periodicity is shown in the studies of Von Voss (4), Flügel (5), Freeman and Wonderlic (6), and others, when the time interval is equally short. The following study was made with the view of investigating phenomena that might emerge when the work is prolonged to 6 or 7 hours.

Subjects.—The Ss of this experiment were 12 College men. They had consented to work continuously at their maximum possible rate till they were too exhausted to continue, and were aware that this might take from 6 to 7 hours at least. Their motivation was excellent and their performance, most praiseworthy, for they continued their task for that lengthy period in spite of boredom, physical pain, and distress. The author wishes to express his appreciation of their efforts.

The Task.—The task was a simple one; it consisted of tapping with a metallic stylus at maximum possible speed, alternately on two brass plates, 4″ square, situated 15″ apart. Any sort of position or movement was permitted. The counting was done by electric counters and the timing taken by a stop watch, checked by an ordinary watch. As the timing was done in 10 sec. intervals during certain phases of the experiment at least 2 counters were connected in the circuit, so that either one could be rapidly thrown in the circuit to permit leisurely reading of the counter. Most of the time an additional counter was used, either as a check, or held in reserve in case of break-

* Paper delivered at the Ottawa meeting of the American Association for the Advancement of Science, June 29, 1938. The work was done at Alberta.

down. Scoring was done by at least 2 E's, who recorded the number of taps per interval, and who were instructed to maintain S's performance at a maximum rate, by warning him when his scores were dropping below a predetermined limit and by occasional verbal incentives.

Brief preliminary practice was given S to accustom him to the experimental procedure, and to permit him to determine the manner of holding the stylus best suited to him. About one week prior to the main experiment, a 10 min. test was given to find S's maximum working rate. In order to insure maximum effort during the main experiment it was decided that should S's score approach 80 percent of the average rate as determined by the 10 min. test, he should be warned and urged to speed up. When he dropped below 80 percent for 3 mins., S was to be allowed a brief rest, which varied from 1–10 mins. No other rests were permitted, except those due to failure of the apparatus. No S had more than 4 rests, some had none, and the average was 2.5 rests with a mean time of 7 mins. After a few Ss had completed their main experiment it was seen that 80 percent was too low a minimum rate.[1] Accordingly the minimum set for the remaining Ss was the average rate made by each S during the pre-experimental period. Some Ss easily exceeded this minimum.

Results

As may be seen from Table I and from the curves, surprisingly little work decrement was shown in general even after several hours of work. Ss are able to maintain a very high

TABLE I

Efficiency of Performance

The rates are the number of taps per min. Times are given in minutes.

Ss	Pre-Exper. Rate	Test Rate	Total Time	Total No. of Taps	Range	σ Dist.	Time of Max.
1	297.0	278.7	345	66, 152	226–346	18.0	25
2	244.0	255.8	355	83, 135	232–279	24.0	55
3	289.4	274.5	430	118, 035	245–299	15.6	165
4	216.6	215.9	420	90, 661	190–243	19.0	115
5	232.5	248.3	390	96, 873	203–275	23.5	140
6	196.4	214.8	420	90, 195	198–235	13.2	205
7	214.7	261.5	355	92, 833	227–295	22.9	230
8	238.0	277.3	426	118, 143	243–343	30.9	40
9	246.5	245.7	355	87, 213	190–286	24.0	75
10	260.5	214.3	275	58, 933	170–244	15.6	85
11	249.0	251.1	360	90, 410	190–278	18.7	40
12	263.5	260.0	355	86, 300	241–294	25.1	325

output of work for long periods, when motivation is sufficiently great. This work output is not constant, however, and is subject to considerable fluctuations. The data were plotted

[1] From Table I it may be seen that the test rate was greater than the pre-experimental rate for 5 Ss; the same, for 4 Ss; and less, for 3 Ss.

for average tapping rate per min. against time. The tapping was estimated for some Ss by 1, 3, 5, and 10 min. intervals. The 5 min. interval was selected as it seemed most effective in eliminating random fluctuations without hiding the main trends. Accordingly the data were estimated on the average tapping rate per min. over a 5 min. interval.

Periodicity

When the data were plotted the periodicity of the curves was at once apparent and it was evident that no parabolic, hyperbolic, nor logarithmic equation could be fitted to them. As an inspectional method of estimating periodicity is too subjective, it was determined to apply statistical criteria. Periodicity in output has been studied from time to time, but no statistical test which has yet been applied seems thoroughly satisfactory. Workers have used, at various times, modal or bimodal distribution of incidence of maxima and minima (**6**), tetrachoric correlations of phase departures from smoothed means (**7**), tabulation of deviations exceeding 1 A.D. from smoothed curves (**6**), and, most frequently, mere inspection of the work curves (**4, 8, 9, 10**). Since meteorologists must frequently evaluate the periodicity of their data it was felt that their procedures might be helpful. Accordingly the procedures outlined by Marvin and Clough were applied to the data.

Clough (**11**) reviews several tests for periodicity and gives as the most significant criterion one which he devised based on smoothing formulæ. He tabulates the frequency of the intervals from maximum to the next maximum, and from minimum to the next minimum, in a series of 900 unrelated numbers smoothed by (*a*) $\frac{a+b+c}{4}$ and (*b*) $\frac{a+2b+3c+2d+e.}{9}$ The corresponding values obtained from the data which are being tested for periodicity are compared with those obtained from the random series. Marked divergence of most frequently occurring intervals in the two series is taken to indicate periodicity, in which, of course, the phase is not necessarily constant. As the judging of the divergence of the most

frequently occurring intervals in the two series may become more subjective than the inspection of periodicity from the curves themselves, a modification of Clough's method was used. It simply consists in calculating the mean interval for each series and estimating periodicity when the mean intervals of the data being tested are definitely greater than the corresponding ones of the random series. This method has the advantage of summarizing the data more succinctly and at the same time of bringing out more clearly dissimilarities in the two series. The comparison is shown in Table II. The

TABLE II

CLOUGH'S CRITERION FOR PERIODICITY

Ss	Unsmoothed Data	Data Smoothe by $\frac{a+b+c}{4}$	Data Smoothed by $\frac{a+2b+3c+2d+e}{9}$
Random Series	3.00	4.53	6.85
S 1	3.78	6.20	9.27
S 2	3.33	4.89	9.26
S 3	3.63	5.98	12.00
S 4	3.60	5.79	8.01
S 5	3.19	5.26	9.73
S 6	3.73	5.85	8.70
S 7	4.21	4.49	7.74
S 8	3.93	5.43	8.31
S 9	3.28	6.12	7.40
S 10	3.46	5.54	8.50
S 11	4.14	5.89	7.21
S 12	3.22	4.29	6.80

periodicity of only one S (No. 12) is doubtful; all the other Ss show pronounced periodicity.

Marvin calls his method the periodocrite and describes it as follows (12):

"Tabulate the data in rows and columns in a manner designed to bring into the same columns data of the same suspected phase relations. In carrying out a tabulation in rows and in columns it is equally possible to place in each of the several columns the same phase values of the data when the length of the period is variable as when it is constant.

Let n = the number of cycles,
p = the number of phase values.

Let $\sigma_{np} = \sigma_o = \sigma$ of all the data,

$\sigma_n = \sigma$ of the p mean phase values of a tabulation.

Then $\frac{\sigma_n}{\sigma_o} = \frac{1}{\sqrt{n}}$. (1)

Let y = ratio $\frac{\sigma_n}{\sigma_o}$ which we may regard as a coefficient of variation.

Also let $\frac{1}{\sqrt{n}} = x$.

Hence from (1) $y = x$.

Such a line represents the results of a perfectly fortuitous combination of data employed.

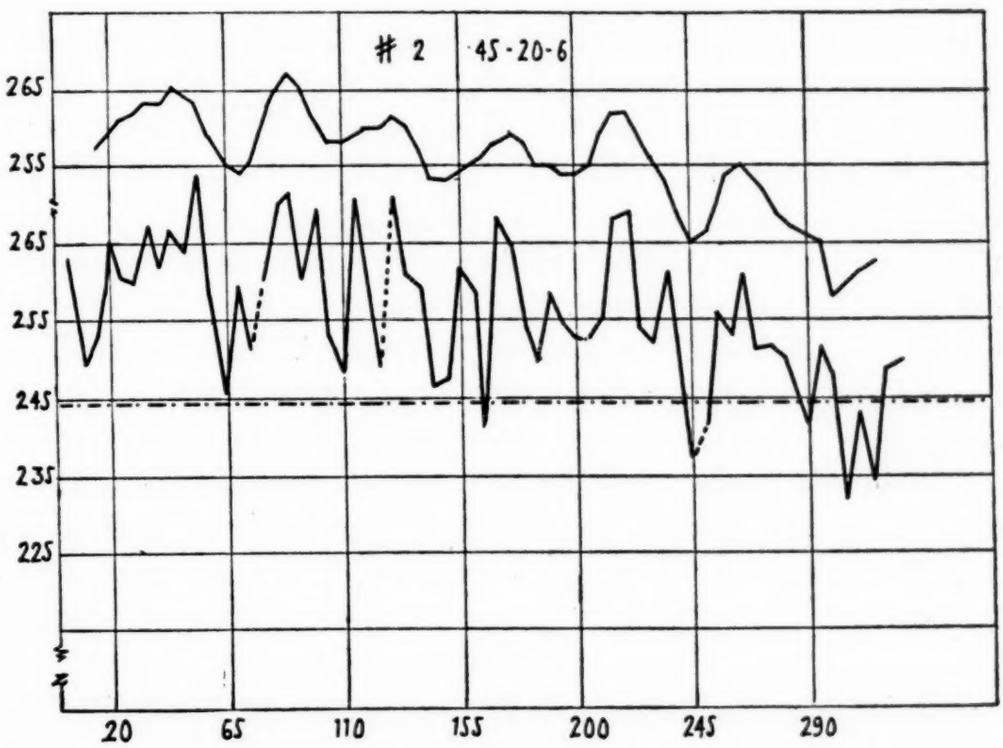


FIG. 1. Curve of work of S, No. 2, plotted in 5 min. intervals. Abscissæ—time in mins.; ordinates—taps per min. Dashed line is the pre-experimental rate (244.0 taps per min.). Dotted lines indicate rest periods. Vertical guide lines placed at the beginning of each cycle. Upper curve is the data smoothed by $\frac{a + 2b + 3c + 2d + e}{9}$. Lower curve is the unsmoothed data, and the ordinates apply to it. Legend 45-20-6 indicates that the length of the period is 45 mins.; that the starting of the first cycle is 20 mins. from the beginning; and that there are 6 cycles in all.

"From the foregoing we see in general that to test a periodicity it suffices to form one or more tabulations as explained and compute values of x and y for one or more groups of data. When y is substantially and consistently greater than x a real periodicity is indicated of greater or less amplitude. If x and y are nearly equal, especially y smaller than x, the amplitude of the periodic variations is less or no greater than that wholly due to chance. In the face of such a result the probability of the cycle being real is very small or nil.

"It will be noticed that theoretically y cannot exceed unity. However, when n is large y may become several times larger than x. This shows strongly marked periodicity but in all such cases inspection alone establishes the same fact, and the calculation of x and y simply serves to express in mathematical terms the proportionate effects of periodic control as compared with chance as represented by x."

It is to be noted that Marvin's periodocrite is tantamount to determining the reliability of the curvilinear correlation ration (η) between the average phase values of the cycle and the mean phase angle. Blakeman's criterion for linearity $[N(\eta^2 - r^2) > 11.37]$ may also be used to give a precise measure of periodicity, for evidently the degree that the curvilinear correlation ratio η differs from the rectilinear correlation coefficient r is a measure of the amount of periodic oscillation.

Marvin's periodocrite and the criterion of Blakeman have accordingly been used to determine regularity of periodicity and the results are shown in Table III.[2]

Results

Definite periodicity sufficiently regular to fulfil all these criteria is shown by 7 of the 12 Ss. For 2 of these Ss, however, slight adjustments had to be made in the periodic tabulation. In one case (S, No. 6), the period was lengthened from 55 mins. to 60 mins. about midway through the test. Another S (No. 5), probably due to a pronounced decrease in motivation, manifested a marked work decrement for about

[2] Schuster's periodogram method was also tried on a few curves and did not seem to yield better results, although more unwieldy and time consuming (13).

TABLE III

CRITERIA OF PERIODICITY

Col. 1—S's identification number.
Col. 2—Length of the period in mins.
Col. 3—Time in mins. from start of work to beginning of first cycle.
Col. 4—Number of cycles shown.
Col. 5—The curvilinear correlation ratio (η).
Col. 6—Marvin's periodocrite; for periodicity it should exceed 1.00.
Col. 7—The rectilinear correlation coefficient (r).
Col. 8—Test for linearity of regression, and its reliability $\left(\frac{\zeta}{\sigma_\zeta}\right)$, Col. 9.
Col. 10—Blakeman's criterion; it should exceed 11.37.

(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
1	40	15	8	0.952	2.69	−0.089	0.899	3.81	57.5
2	45	20	6	0.860	2.13	0.277	0.663	2.99	35.9
3	60	10	6	0.638	1.56	−0.087	0.399	2.17	28.7
4	65	10	6	0.570	1.40	0.017	0.325	2.52	25.4
5	70	30	5	0.593	1.33	0.165	0.324	2.35	22.7
6	55	15	3	0.452	1.20	0.035	0.204	2.04	16.5
	60	175	4						
7	60	40	6	0.453	1.11	−0.020	0.206	1.93	14.8
8	45	30	9	0.318	0.96	0.000	0.101	1.37	8.2
9	65	35	5	0.383	0.86	0.230	0.093	1.23	6.0
10	35	30	6	0.388	0.95	0.195	0.114	1.10	4.8
11	40	45	7	0.307	0.81	−0.106	0.083	1.11	4.7
12	25	0	14	0.204	0.80	0.000	0.042	0.87	3.0

25 mins. at the beginning of the last third of his test. When this 25 min. interval is omitted the rest of the data satisfies the criteria. In spite of several periodic tabulations, 5 Ss did not satisfy the criteria. They did however show marked oscillations in their output and undoubtedly on the basis of the Clough criterion have an irregular periodicity.

From Table III, Columns 2, 3 and 4, it may be seen that on the average there are 7 cycles, with an approximate length of 50 mins. per cycle. Moreover the period does not start till after about 25 mins. from the beginning of the work period, as the initial spurt, initial decrement, and adaptation are so pronounced as to mask the periodic character of the work curve.

It is to be noted that owing to the extraordinary motivation shown by all Ss, the output, both as to amount and to rate, far exceeds that of normal work, and hence the results

apply only to the special conditions obtaining during this experiment.

Discussion

That periodicity should be found in work curves where variability is a commonplace is to be expected. The difficulty, nay impossibility, of S maintaining for any length of time a pre-selected efficiency, is such that there is a continuous variation in working rate from second to second and even from tap to tap, as was shown by Bliss (**14**). The natural rhythms of the body and the well known oscillations of attention suggest that these fluctuations of efficiency might tend to group themselves periodically.

In any study of continuous work there are present two opposing groups of conditions, generally called practice and fatigue. Tentatively one may define practice as any combination of factors tending to improve the performance; and fatigue, as any combination of factors that has the contrary effect. Specifically, practice consists, in part at least, of the following factors:

(*a*) Initial adjustment and adaptation to test conditions.
(*b*) Elimination of useless and detrimental movements.
(*c*) Acquisition of more effective movements, which are generally worked into some kind of suitable rhythm.
(*d*) The shifting of the attention from a foveal type, which is too sensitive to distractions, to a peripheral type, where the operations tend to become semi-automatic.

The concept of fatigue is one of baffling complexity, but the following components may be noted:

(*a*) On the physiological side there are depletion of muscular energy, formation of toxins, inadequate circulation, etc.
(*b*) Muscular cramping and mental blocking, involving loss of rhythm.
(*c*) Intrusion of conscious factors, and sensory distractions.
(*d*) Loss of incentives and discouragement.

These factors are constantly operating, but not always to the same degree. Practice and fatigue have opposite effects

on efficiency, and the equilibrium between the two sets of forces is quite unstable. As stated by Flügel (**5**): "Practice is nowhere rigidly determined by any quasi-mechanical 'limits' but in its actual occurrence is (like fatigue) always a matter of the relation between the energy available and the obstacles to be overcome." Were this relationship a simple one, and not subject to upsets in a very unstable equilibrium, there would be close adherence to the parabolic and hyperbolic types of work curves of Ioteyko (**1**) and others, with practice being dominant at the outset of the work and fatigue at the latter part of the work, provided of course it is carried on for a sufficient time.

One should not place too great a reliance on this simple explanation of Ioteyko, even though she gives a plausible interpretation in physiological terms to the constants of her parabolic equation, since Neifeld and Poffenberger (**2**) found contrary signs should be placed before the equivalent terms in their equations of the work curves. They state: "These discrepancies among individuals and among the records of the same individual make us hesitate to adapt any sweeping generalizations from them concerning the nature of fatigue. The equations must be recognized as only trends that happen to fit the observed data."

Though any explanation of the complicated processes that take place in a long period of work must necessarily be inadequate, the following tentative account of the factors that cause periodicity is here advanced as a first approximation to the actual happenings. It is assumed that maximum possible efficiency is rarely if ever attained and then only for the briefest time, when it is manifested by sudden peaks of the curve under intense effort. Often the highest peaks occur quite late in the work curve, as may be seen from Table I, when S must be quite fatigued, so that even the fastest rate at these points does not represent his physiological limit. In general the work is carried on at a much lower level, S's effective maximum. At this level the equilibrium between practice and fatigue is of a more stable type, and tends to yield in general the parabolic curve of Ioteyko, in which a

steady rise in efficiency is followed by a slower prolonged drop. However, at the outset of the work period, adaptation to working conditions, initial spurt caused by S's freshness and eagerness, and the initial decrement, mask the regularity of the curve and delay the appearance of periodicity.

Later the semi-automatic stage appears (in our results not sooner on the average than 25 mins. from the start of the work), at which periodic oscillations may emerge. They are caused by grouping of sets of factors that tend to work in the same direction. Thus, for instance, a decline in the work curve, which comes on rather abruptly after a steady period at the effective maximum, may be caused by temporary boredom, or decreased motivation; conversely, too labored an effort, with consequent intrusion of conscious factors, blocking, or cramping of muscles, may also result in decreased efficiency, which, as Bills (**15, 16, 17**) has pointed out, gives short pseudo-rest periods, and acts as Nature's preventative of complete exhaustion. Then too, invariably after a spurt immediately follows a compensatory drop of longer duration. In this regard Flügel (**5**) states: "When working at great pressure an individual's efficiency seems to remain for the most part at a level near his maximum attainment and is broken, not so much by brief spurts of unusual brilliance, but rather by periodic relaxations, each resulting in a short, but well marked decline in output." S is generally aware of his poorer performance and attempts to regain his effective level. His motivation is increased, consciously or mechanically he regains his rhythm, often after making purposely awkward, unnecessary movements or shifting his position, his efficiency slowly picks up, and he may and frequently does exceed his effective working rate. But this involves the expenditure of more than his normal energy output, and the negative factors once more become predominant, the fall in efficiency is rapid, and the cycle repeats itself.

These sets of factors tend to group themselves around natural rhythms, peculiar to each individual. Von Voss (**4**) ties this personal variability up with the oscillations of attention and mentions the work of Eckener, Pace, Marbe and

Lehman in this connection. He refers however to periodicity of much shorter duration than appears in this paper.

Periodicity is contingent upon the maintenance of a high degree of motivation, with the consequent cutting down of extreme variability. As has been pointed out by Freeman and Wonderlic (**6**), Ss who manifest great variability do not show periodicity. Discouragements, distractions, pain, social disturbances, all tend to disturb the rhythm and to mask periodicity. Often they are violently opposed by forced spurts, which in turn are followed by the usual decrements. The oscillations that result are not of a regular character, and give a work curve that does not conform to the criteria here proposed for periodicity.

This rhythmic process may be carried on with unusually high efficiency for much longer periods than normally would be imagined, because there is a marked degree of economy of effort, and the individual profits by every chance of minor relaxation that tends to ward off ultimate exhaustion. It is to be noted also that the encouragement which S gets from the success of his efforts piles up extra incentive that greatly defers his apparent exhaustion. Even then, towards the end of the work curve, the increase of variability with fatigue is due rather to a loss of control than directly to the exhaustion of muscle (*cf.* Weinland (**18**)). Therefore this study does not show anything like the complete exhaustion of ergographic studies; there is, moreover, no possibility of forcing maximum efficiency by exterior control other than has been mentioned, and the individual is largely his own control.

(Manuscript received November 1, 1938)

References

1. Ioteyko, J., *Les lois de l'Ergographie*, 1904.
2. Neifeld, M. R. and Poffenberger, A. T., A mathematical analysis of work curves, *J. Gen. Psych.*, I, 448–458.
3. Philpott, S. J. F., Meaning of G and other factors in terms of rates of output, *Brit. J. Psych.*, 1936–7, **27**, 225.
4. Von Voss, G., Über die Schwankungen der geistigen Arbeitsleistungen, *Psych. Arbeit.*, 1899, **2**, 399–450.
5. Flügel, J. C., Practice, fatigue and oscillation, *Brit. J. Psych., Mono. Supp.*, **13**, 1–80.

6. Freeman, G. L. and Wonderlic, E. F., Periodicity in performance, *Amer. J. Psych.*, 1935, **47**, 149–151.
7. Chen, L., Periodicity in oscillation, *Brit. J. Psych.*, 1934–5, **25**, 383–392.
8. Philpott, S. J. F., Theoretical curve of fluctuations of attention, *Brit. J. Psych.*, 1934–5, **25**, 221–256.
9. Entwhistle, W. H., Oscillation, *Brit. J. Psych.*, 1936–7, **27**, 313–320.
10. Wheeler, R. H., *The Science of Psychology*, 1929, p. 306.
11. Clough, H. W., A statistical comparison of meteorological data with data of random occurrence, *Month. Weather Rev.*, 1921, **49**, No. 3.
12. Marvin, C. F., Theory and use of the periodocrite, *Month. Weather Rev.*, 1921, **49**, No. 3.
13. Rietz, H. L., *Handbook of Mathematical Statistics*, 1924, pp. 170–176.
14. Bliss, C. B., Investigations in reaction time and attention, *Yale Psych. St.*, 1893, 1–56.
15. Bills, A. G., Inhibition and facilitation, *Psych. Bull.*, 1927, **24**, 473–487.
16. Bills, A. G., The influence of muscular tension on metal work, *Amer. J. Psych.*, 1927, **38**, 227–251.
17. Bills, A. G., Blocking; a new principle of mental fatigue, *Amer. J. Psych.*, 1931, **43**, 230–245.
18. Weinland, J. D., Variability of performance in the curve of work, *Arch. Psych.*, No. 87, pp. 69.

PERCEPTUAL FLUCTUATION AS A FATIGUE INDEX

BY H. L. HOLLINGWORTH

Barnard College, Columbia University

Among the proposed indices of general fatigue is the rate of fluctuation of various plane figures which are perceived in equivocal perspective. McDougall,[1] for example, has referred to this as 'a very delicate method' which he 'devised for revealing the influence of alcohol and other drugs on the brain.' The method is described by this writer:

> It consists in setting the subject to register the successive alternations that occur when he continuously observes an ambiguous object, such as any one of the well-known 'illusions of reversible perspective,' the transparent cube, the open book, or the stair-case figure.

McDougall describes the wind-mill figure which he prefers and the intervals at which it reverses its apparent direction of rotation, and adds:

> The duration of these intervals varies greatly from one subject to another, and is characteristic for each subject in his normal state. Alcohol, ether and chloroform produce very marked prolongation of these intervals in all subjects tested. General fatigue has a similar though less pronounced effect, slowing the rate of alternation. The alkaloids, strychnine, tea, and coffee (also opium, although here my observations are very few) produce the opposite effect; namely, an increase of the rate of alternation.

On the basis of this strong recommendation two such figures were used in a series of experiments on the possible effects, on psychophysical performance, of such rhythmical motor activities as those involved in chewing, while working.

The results were so contradictory to those reported by McDougall that it appears desirable to record them. They also have a certain intrinsic interest, as well as an interesting relation to the original problem of the experiments on the influence of mastication.

In these experiments a group of six subjects worked

[1] McDougall, Wm., *Outline of Abnormal Psychology*, p. 74.

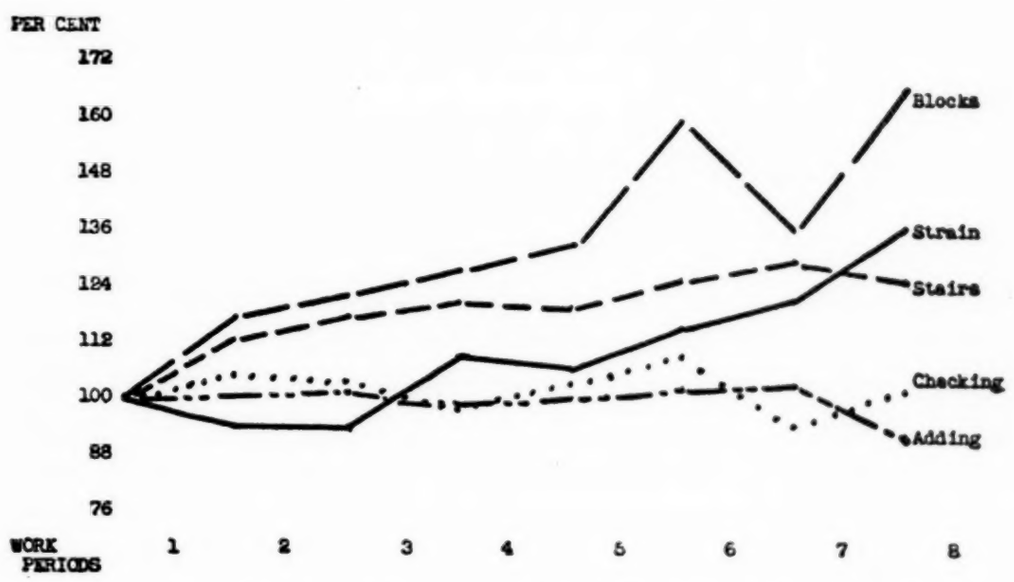
steadily at various types of work, passing regularly from one type to another until a unitary series of tasks, requiring about 40 minutes, had been accomplished. The whole series was then repeated, eight times without other interruption than 30 minutes for a light lunch which was eaten in the laboratory. This regime was continued for six successive days. These tasks were strenuous, the work program rigorous in its demands; and the subjects were strongly motivated toward effort. If there is such a thing as 'general fatigue' it appears wholly reasonable that these workers experienced it on all six of these experimental days.

Two kinds of indicators of fatigue effects are present, in addition to the indications from the equivocal figures. One is the rate of work, with prescribed accuracy, in such activities as number checking and simple addition. The work curves for these two activities here given are for 3-minute tests of number checking (Woodworth-Wells) and 5-minute periods of simple addition (adding the digits in the groups on the number-checking form). The total work by all six subjects, on the five days after the first preliminary practice day, is used as the measure, stated as percent of the work of the first hour, for each day.

The curves do not show any considerable work decrement, although it is true that the poorest work in each case is done in the later hours of the day, and that variability from trial to trial appears to increase toward the end of the day. However, as the days proceeded, considerable practice gain accrued. In addition the total number of problems worked at a sitting increased from 264 to 350. In checking the number groups correctly discriminated increased from 899 to 1130. Throughout the six days, that is to say, gain was being made. But gain does not appear within the day's work, as shown by these curves. The inference seems justified that such work decrement as the curves do show occurs in the face of a steady trend toward improvement. To do no better than maintain a constant level of accomplishment would under such circumstances be a sign of fatigue.

After each round of tasks the workers rated their feelings during the round in question, on a linear rating scale, ranging from Extreme Comfort, Ease and Relaxation at one end to Extreme Strain, Effort and Discomfort at the other. We have considered the Relaxation end as zero and the Strain end as 20, and thus tried to give quantitative form to these subjective indications of the workers' affective experience while working. As shown on the graph, after a brief warming up period the curve moves steadily upward toward Strain. Even when the work done maintained a fairly constant level, it was accomplished with greater subjective difficulty. Perhaps this is as good an indicator of fatigue as any now available. During every round of work (test series) the subjects observed for two minutes each two of the familiar illusions of reversible perspective. One was the ordinary stair-case, drawn to large size so as to be clearly visible from a distance of fifteen feet. The other was the familiar nest of cubes, which changes in apparent arrangement and number, also drawn to large size. Each subject held in his hand a Veedor counter on which was recorded each change in the appearance of the figure. The number of alternations in two minutes thus becomes the measure used for each of the two figures.

Curves for each figure are shown on the graph. From the



Alternation of equivocal figures, strain, and diurnal variations in work.

beginning of the day both figures increase in rate of fluctuation, the increase amounting to about 25 percent in the case of the stair-case and 50 percent in the case of the blocks. These curves, in fact, move upward quite parallel to the Strain curve.

If, therefore, the work curves show genuine fatigue, and if the subjective ratings of strain and relaxation are also a valid indication of the fatigue condition, we must conclude that the effect of fatigue on the alternation of the equivocal figures is to *increase* the rate of fluctuation, thus *shortening* the intervals of appearance, rather than prolonging them, as reported by McDougall. The fatigue effect appears to resemble that which McDougall attributes to the alkaloids, rather than that which he cites for alcohol, ether and chloroform. This being the case, check experiments on the reported effect of the various drugs would also appear to be in order.

Check Experiments

Following up the suggestions from the foregoing results, further experiments were performed with four adult subjects, working four days, with six half-hour work periods both forenoon and afternoon. In each half hour the block figure (Scripture's) was observed for one minute and the spontaneous fluctuations recorded by each subject. The totals for all four subjects are as follows:

	Forenoon Work Periods					
	1	2	3	4	5	6
May 18.....	62	59	65	67	61	65
May 19.....	52	61	56	62	67	55
May 25.....	60	50	55	59	48	53
May 26.....	59	59	56	57	58	49
Total in 16 min.....	233	229	232	245	234	222

First 3 trials—694 Last 3 trials—701

Total forenoon—1395

	Afternoon Work Periods					
	1	2	3	4	5	6
May 18.....	59	55	50	71	58	58
May 19.....	50	56	59	44	51	56
May 25.....	43	53	55	55	47	41
May 26.....	54	54	51	39	51	56
Total in 16 min.....	206	218	215	209	207	211

First 3 trials—639 Last 3 trials—627

Total afternoon—1266

There are about 10 percent fewer changes in the afternoon than in the forenoon; but throughout neither period is there any consistent change in the rate of fluctuation as a function of continued work. Since a rest and lunch period intervened, there is no reason for assuming that the afternoon was a more 'fatigued' period than the forenoon. The drop in rate of fluctuations appears in the very first afternoon trial, and is merely maintained. Hence we can not use this comparison as evidence of slower fluctuations with greater fatigue.

Along with the observations of the changing perspectives, certain other kinds of work were measured—tapping for 1 minute, 3-hole coördination for 1 minute, and cancelling A's for 3 minutes. If we take the scores made in the last three trials of each of the work periods and compare them with the scores made in the first three trials each time, this will give us a reasonable index of out-put change. The ratio of the last half to the first half of the work period will show whether work was, relatively, improving or falling off.

The following Table shows these indices for each work period. They were first computed for each worker, then averaged to get the figures given here for the four workers combined. On the whole, the work of the latter half is better than that of the first half of a period; but the fluctuations tend if anything to be *slower* in the second half—not when out-put shows signs of fatigue, but when work scores are relatively better. Again the evidence is against slower fluctuation as a sign of fatigue—although out-put is by no means a satisfactory index of fatigue.

RATIOS OF LAST THREE TO FIRST THREE TRIALS OF HALF-HOUR WORK PERIODS, 4 WORKERS COMBINED

	Four Forenoons	Four Afternoons	Character
Fluctuations per minute	102.5 91.5 98.2 95.8	112.0 114.5 97.5 98.4	3 faster 5 slower
Coördination hits per minute	98.8 104.0 101.5 101.2	105.0 103.2 99.1 104.7	2 poorer 6 better
Cancelling A's per 3 minutes	105.9 99.7 104.5 108.7	103.0 99.7 97.8 101.1	3 poorer 5 better
Tapping hits per minute	99.6 97.5 99.0 97.9	99.5 102.7 100.7 102.9	5 poorer 3 better
Feelings of strain for half-hour	105.0 105.0 93.5 94.0	98.0 118.0 107.0 93.5	4 worse 4 better

As a matter of fact there is but slight indication (5 periods out of 8) that there is any change in the rate of fluctuation with this group of subjects. That there should not be any considerable change is suggested by the fact that on the whole there was no general increase in subjectively estimated feelings of strain (4 cases out of 8). The two negative results give a wholly consistent picture.

In each half-hour period the subjects rated their feelings of strain and relaxation on a rating scale; higher scores on this scale mean greater degrees of felt strain and tension. In at least one sense, this is an indication of degree of fatigue—subjective fatigue feelings. We here arranged the eight half-day periods in rank order, putting first that period in which there was the most conspicuous change in fluctuations (increase) in the second half as compared with the first half. Alongside are similar measures of the relative tension and strain scores, a higher figure again meaning that this factor was greater (more strain) in the second half than in the first half of the work period.

	Fluctuations	Felt Strain	
	114.5	118.0	
Highest	112.0	98.0	Average
four	102.5	105.0	107.0
	98.4	107.0	
	98.2	93.5	
Lowest	97.5	93.5	Average
four	95.8	94.0	96.5
	91.5	105.0	

The fluctuations of the reversible figures are relatively more rapid when the felt strain is relatively greater. In so far, at least, as subjective feelings of strain are any indication of relative fatigue, rapid fluctuations, not slower ones, go along with greater fatigue.

Continued work does not necessarily involve fatigue. Work out-put does not necessarily reflect fatigue, particularly if, as in the present instance, practice is also a factor. Of course fatigue may be *defined* so as to mean simply the effect of continued work, or the presence of reduced out-put. But we may also mean by fatigue a feeling which the worker has; and as thus defined, the fatigue experience is accompanied by *more rapid*, not by slower, fluctuations of figures of reversible perspective.

These experiments are reminiscent of those of Ash [2] who practiced his observers in deliberately producing the alternation of several of these reversible perspectives. Ash found that these voluntary reversals were slower after continuous work than before it, and interpreted this to be a lack of 'control' and a sign of 'mental fatigue.'

But Ash also did a few experiments with 'spontaneous reversals,' although on the same observers who had been trained to effect voluntary alternation in the figures. He makes no use of these data except to state that they are 'in part' similar to the data for deliberate reversal, but fortunately he gives the records for each subject.

He used two work periods, forenoon and afternoon, separated by a two-hour rest period, with four trials in the forenoon and five in the afternoon. With the block figure

[2] Ash, I. E., Fatigue and its effects upon control, *Arch. Psychol.*, No. 31, June, 1914.

which he used, seven of the nine sets of results show *more rapid* alternation in the later hours of the work period. With the stair-case figure, five of the nine complete sets of results showed the same tendency. The averages, which I have computed from his table of records, show a clear enough tendency toward more rapid spontaneous fluctuation, even in observers previously practiced in control, and whose voluntary reversals were slower, after continued work. These results do not agree with McDougall's report, and are in complete harmony with our own observations.

Ash, furthermore, devotes a page to expressing his discontent with the block figure (Scripture's form) which he used, finding the results of its use 'less uniform than with the others.' With this figure his observers could not easily find a technique for voluntarily producing the reversals. "They came rather fortuitously." In the few experiments with spontaneous reversal which he reports it is just the 'unsatisfactory' figure which most clearly shows the increased rate of fluctuation in the later hours of work. This is another bit of evidence suggesting that when the alternations recorded are those that 'spontaneously' occur, fatigue, if it does anything, increases the rate of fluctuation and *shortens* rather than *prolongs* the intervals.

Most interesting of the results given by Ash are two tables of data showing the rate of reversals for several observers and three different figures, at successive hours of the day from 7 A.M. to 10 P.M. "All of them did more or less work during the day," but since they did not do Ash's work in *Addition* the results are called by him 'the effects of practice'! He found that "There is a decided increase in the rate of reversals," of both the voluntary and the spontaneous variety. It does not seem unlikely that it was even more 'fatiguing' for these observers to add to their regular work these numerous expeditions to the laboratory than for them to settle themselves there for a period of continuous simple addition. The increased rate of fluctuation of the reversible figures points strongly in this direction.

Ash was quite right, however, in his surmise that practice

increases the rate of spontaneous fluctuations. Just as the level of work in Adding and Checking, in our own records, rises from day to day, although the course of work within a given day is fairly uniform (without marked work decrement or increase), so also the rate levels of these fluctuations rise in successive days of work and this is perhaps the clearest indication of the relatively permanent effects of practice.

We conclude that the rate of fluctuation of illusions of reversible perspective may very well afford a neat index of fatigue, but that the present evidence is that fatigue *increases* rather than retards the rate of alternation of such figures, when this alternation is spontaneous.

We believe also that the case of voluntary alternation merits further investigation, and that conclusions concerning the effects of various drugs on these changes may be subject to revision. Furthermore, better explanations of the phenomena are needed—particularly, better than current demonological interpretations in terms of 'field forces' and the ethical aspirations of 'gestalten.'

(Manuscript received October 14, 1938)

THE KNEE-JERK AS A MEASURE OF MUSCULAR TENSION *

BY FREDERICK A. COURTS

University of Missouri

Jacobson and Carlson (3), and Jacobson (4) have found that during relaxation the knee-jerk is greatly reduced in amplitude, suggesting that definite relationships exist between this response and voluntary control of muscles. Diminution or absence of the knee-jerk during sleep has been demonstrated by Lee and Kleitman (6). Among clinicians it has long been common knowledge that the knee-jerk may be reinforced by having the patient clasp his hands and pull, maintaining a strong muscular contraction while the stimulus is being presented. That the knee-jerk is facilitated by attention to and voluntary contraction of muscles other than those responding to the tendon-tap has been observed by numerous investigators since the early study by Lombard (7), which showed that the amplitude of response was increased by environmental influences which were thought to increase the activity of the central nervous system. Lombard's results have been confirmed by Tuttle (8).

As evidence of the generality of the tension that is measured by the knee-jerk, the work of Golla and Antonovitch (2) may be cited. They showed that the amplitude of response is correlated with tonus of the muscle groups producing extension of the knee-joint, and those producing extension and dorsiflexion of the wrist-joint. These investigations suggest that the knee-jerk is facilitated by an increase in muscular tension occurring in various portions of the body.

Dodge and Travis (1) found that photographic recordings

* This investigation was performed in the Psychological Laboratory of Stanford University. The writer is indebted to Dr. E. R. Hilgard and Dr. C. P. Stone for their advice and assistance during the course of the experiment and in the preparation of the manuscript.

of a light wooden lever resting over the quadriceps muscle bear a linear relationship to the amount of tonus induced in that muscle during stimulation of the patellar tendon.

The evidence cited above suggests that since the knee-jerk is reinforced by sustained contraction of various muscles, and since the amplitude of response is correlated with the degree of contraction of the quadriceps muscle involved in the response, the amplitude of the knee-jerk may offer a useful measure of general bodily tonus. The purpose of the present investigation is to study the relationship of the amplitude of the knee-jerk to degrees of tension experimentally induced in musculatures other than those involved immediately in the response to the tendon-tap.

Apparatus

The devices used in stimulating the patellar tendon and recording the amplitude of the knee-jerk are shown in Fig. 1. The stimulus hammer consisted of a one cm metal rod 50 cm in length, which carried an adjustable wedge-shaped piece of wood. The striking edge was 6 cm in width, which permitted a slight amount of side-play to occur without affecting the constancy of the point of contact of the tendon-tap. The hammer was held by an electromagnet at a fixed distance from the subject's knee. When the hammer was released by breaking the circuit leading to the electromagnet, it was pulled rapidly through an arc of approximately 30° by a rubber band 3/8 in in width and 3 in in length. This was fastened to the rod at a distance of 10 cm from the bearing. When the hammer was supported by the magnet, the length of the rubber band was extended to 7.5 in. This arrangement resulted in a light but rapid blow on the tendon. The stimulus hammer was reset by means of a cord running on pulleys to the experimenter's chair. The subject sat in the specially constructed chair shown in Fig. 1, with his leg supported on an adjustable extension arm which was covered with a sponge-rubber pad. His foot was securely fastened so that the lower leg could not swing when the response occurred. An adjustable block on either side of the leg, just above the knee, prevented side movements of the limb.

Muscle thickening was recorded by means of an 8 × 25 mm wooden lever one meter in length, which rested on its edge over the quadriceps muscle approximately half-way between the knee and the crotch. The lever rested on the leg at a point about 24 cm from the fulcrum. A weight of 750 grams was placed on the lever at a point 35 cm from the fulcrum. This served to prevent the lever from bouncing when the response occurred, thus providing a solid base line from which the amplitude of response could be read. The end of the recording lever carried a small metal stylus which rested against a belt of 5-inch Stylographic paper carried by a long-paper kymograph. A screen was placed between the subject and the kymograph.

The movement of the stylus was four times as great as the movement of the recording lever at the point of contact with the leg. Thus, the records obtained with this apparatus were fairly sensitive to slight changes in the amount of muscle thickening in response to the tendon-tap. In reading the records, the length of the vertical excursion of the stylus was measured to the nearest millimeter.

Two procedures were used for inducing muscular tension. One of these required the subject to pull down on a stirrup handle attached to a cable running through an overhead pulley and supporting various weights. The other procedure was to have the subject exert a given amount of pressure on a Smedley hand-dynamometer, which was securely fastened directly in front of the subject. The dynamometer was equipped with a small electrical contact of spring brass, which was connected in series with a small signal light situated on the table in front of the subject. The contact could be set so that when the indicator hand of the dynamometer reached a given point the light would be turned on.

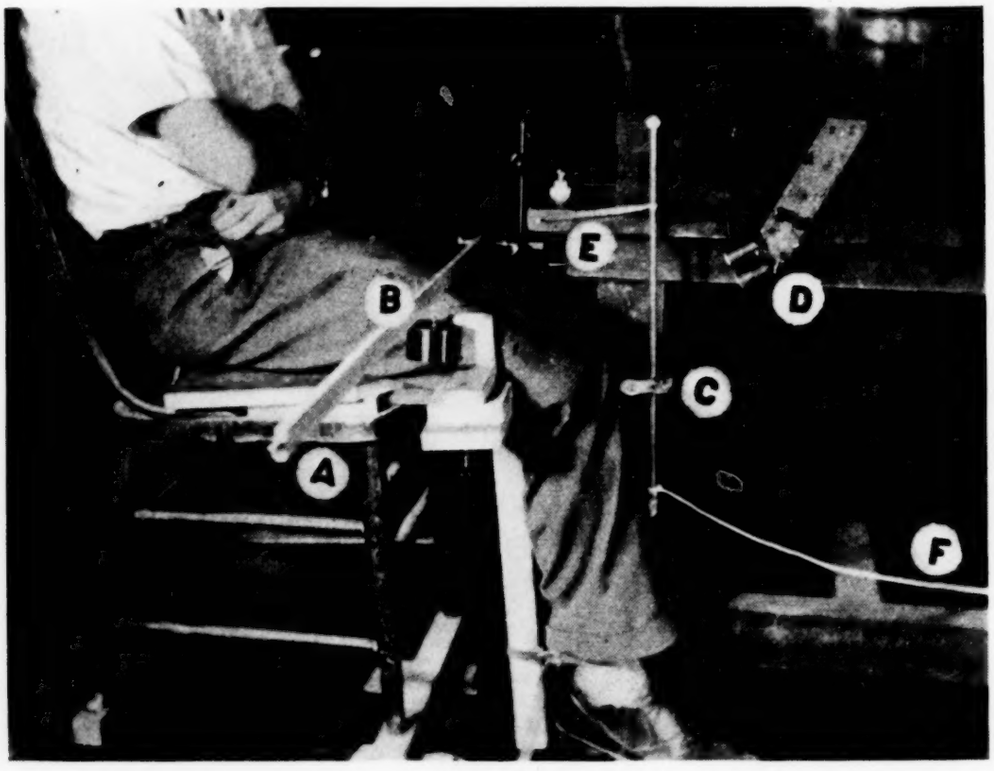


FIG. 1. Apparatus used in stimulating and recording the knee-jerk.

A. Recording lever. *D*. Electromagnet.
B. Weight. *E*. Rubber band.
C. Stimulus hammer. *F*. Cord for resetting stimulus hammer.

EXPERIMENTAL PROCEDURE

Experiment I.—In the first experiment, using 14 male undergraduate psychology students, changes in muscular tension were induced by having the subjects pull down on the stirrup handle. In a preliminary test a spring scales was attached to the cable and the subject was asked to pull as hard as possible for 30 seconds. The reading at the end of that time was taken as an indication of the subject's strength. Twenty knee-jerks were then recorded under each of the following conditions: (*A*) relaxation (S was instructed to imagine himself 'very tired, and ready to fall asleep'); (*B*) normal tension; and (*C*), (*D*), (*E*), (*F*), and (*G*), wherein the subject was supporting weights equal respectively to 1/5, 2/5, 3/5, 4/5, and 5/5 of the reading obtained in the preliminary

test of strength. During preliminary experimentation it was noticed that some subjects tended to anticipate the stimulus when it occurred regularly at intervals of 5 seconds. In order to avoid such anticipatory responses, the tendon-tap was presented in the present experiment at intervals varying at random from 3 to 5 seconds. The order of presentation of the various conditions of tension is shown in Table 1. Ten responses were recorded under each presentation of each condition. After each series of 10 responses, the subject was allowed to rest for 45 seconds while the experimenter adjusted the weights for the next series.

TABLE 1

ORDER OF PRESENTATION OF THE VARIOUS TENSIONS

Subjects	Order													
1 and 8.....	C	G	B	D	F	A	E	E	A	F	D	B	G	C
2 and 9.....	G	B	D	F	A	E	C	C	E	A	F	D	B	G
3 and 10.....	B	D	F	A	E	C	G	G	C	E	A	F	D	B
4 and 11.....	D	F	A	E	C	G	B	B	G	C	E	A	F	D
5 and 12.....	F	A	E	C	G	B	D	D	B	G	C	E	A	F
6 and 13.....	A	E	C	G	B	D	F	F	D	B	G	C	E	A
7 and 14.....	E	C	G	B	D	F	A	A	F	D	B	G	C	E

Experiment II.—With another group of 7 male undergraduate subjects, tension was induced by squeezing the dynamometer with the right hand so as to keep the light on when the contact was set at values corresponding to 1/5, 2/5, 3/5, 4/5, and 5/5 of the dynamometer reading obtained at the end of 30 seconds of maximum effort in a preliminary trial. The conditions were presented in the same order as in Experiment I.

Experiment III.—Another group of 7 subjects participated in this experiment. The procedure differed from that of Experiment II only in that tension was induced by squeezing the dynamometer with the left hand.

Experiment IV.—Sixty undergraduate male subjects participated in this experiment. Tension was induced by squeezing the dynamometer with the right hand. Following a preliminary series of 30 knee-jerks under no experimentally induced tension, the subject was asked to squeeze as hard as possible on the dynamometer handles for 30 seconds. Ten records of the knee-jerk were obtained under each of 6 degrees of dynamometer tension equal to various fractions of the reading at the end of the preliminary test. The conditions of tension were presented in the following order: 1/2, 0, 3/8, 1/8, 3/4, 1/4, 1/4, 3/4, 1/8, 3/8, 0, 1/2. Five knee-jerks were obtained during each presentation of each degree of tension. All subjects followed the same order of presentation.

RESULTS

Figure 2 represents the relationship between the amplitude of quadriceps-thickening of the right leg in response to the tendon-tap and the amount of weight supported by pulling down on the stirrup handle. The 20 determinations of the amplitude of responses which were obtained for each subject under each condition of tension were averaged. Each point on

the curve represents the mean of the averages obtained in this way for the 14 subjects.

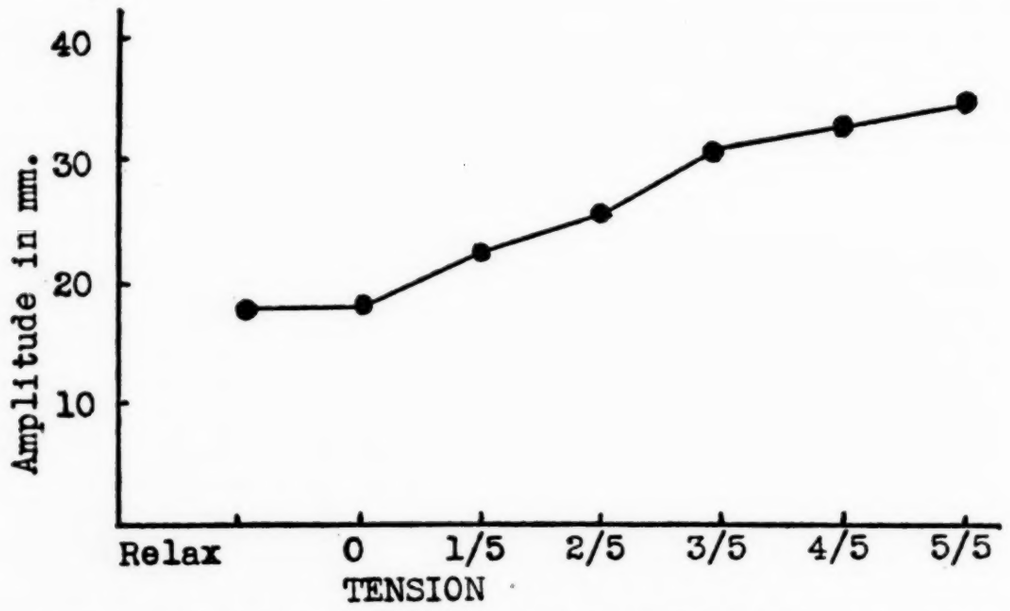


FIG. 2. Relationship of the amplitude of quadriceps-thickening of the right leg to the amount of weight supported by pulling with the right hand (N = 14).

Figure 3 shows the relationship of quadriceps-thickening of the right leg to various dynamometer tensions induced by squeezing with the right hand and with the left hand.

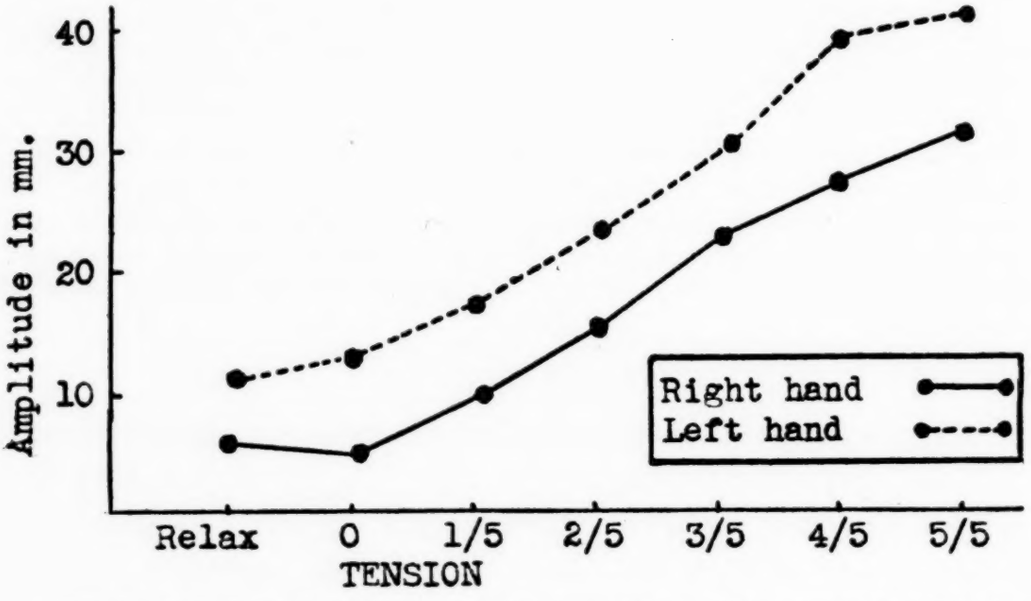


FIG. 3. Relationship of the amplitude of quadriceps-thickening of the right leg to degrees of dynamometer-tension induced by squeezing with the right and left hand (N = 7 in each group).

The values used in plotting these curves were obtained by averaging the individual means, each of which was based on 20 measurements.

The curves in Figs. 2 and 3 show a tendency toward negative acceleration in the region corresponding to extreme degrees of tension. One factor making for such a relationship is the physical impossibility of quadriceps-thickening beyond a certain point. The small difference in amplitude under conditions of relaxation and no dynamometer tension causes the curves as a whole to assume a sigmoid form. That this truly represents the relationship may be doubted because, in this experiment, the condition of relaxation was not adequately controlled. Jacobson (**5**) has shown that the amplitude of the knee-jerk is greatly inhibited by relaxation in the case of subjects who have been thoroughly trained in relaxing. Since our subjects had not received such training, the amplitude of response under relaxation might be expected to be too high due to the inability of some of the subjects to relax. Furthermore, even if the subjects had all succeeded in assuming a comparable state of relaxation, it would be extremely difficult to locate this condition accurately in the series of muscular tensions. The sigmoid character of the curves in Figs. 2 and 3 may be due merely to the fact that the condition of relaxation is not actually comparable with the conditions of experimentally induced tension.

The general similarity of the three curves provides evidence that the amplitude of quadriceps-thickening in response to the tendon-tap may be taken as a measure of changes in muscular tension. The fact that there was a similar change in amplitude with successive degrees of tension induced by means of different techniques, and in different musculatures, suggests that the amplitude of response is influenced by changes in general muscular tension, and not merely by tension induced in a specific musculature. Since the groups are small, differences in the levels of the three curves in Figs. 2 and 3 may be attributed to individual differences in the amplitude of responses under normal conditions of tension.

Figure 4 represents the relationship between the amplitude of response of the right leg and dynamometer tension induced by squeezing with the right hand for a smaller range of tension than was used in the first three experiments. The points on the curve represent the means of the averages of 10 responses obtained from each subject under each condition of tension. The relationship appears to be linear throughout the range of tensions used in Experiment IV. This is essentially what would be expected from the data shown in

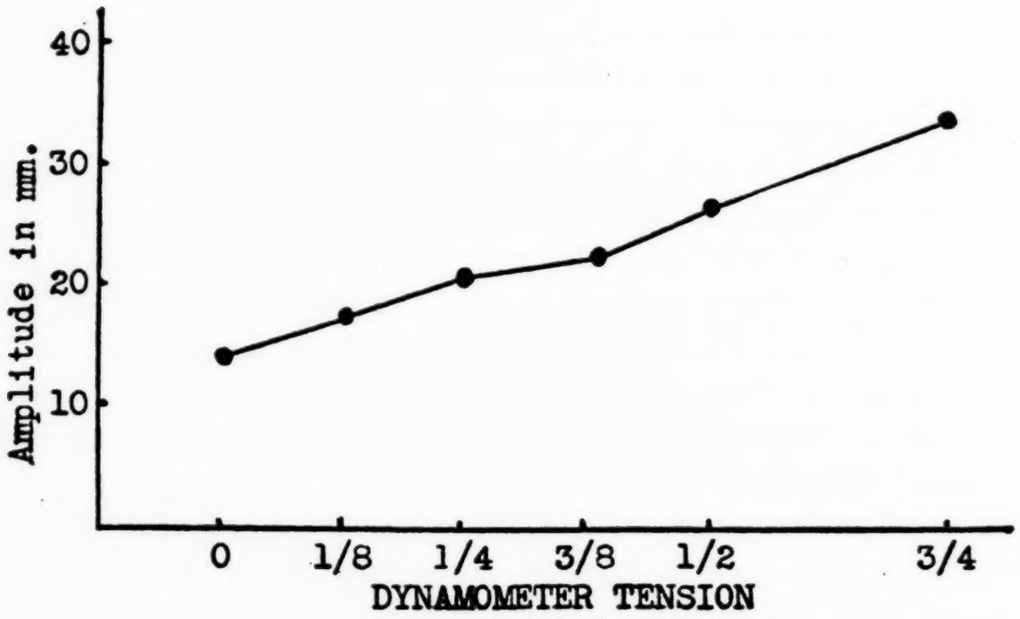


FIG. 4. Amplitude of quadriceps-thickening under various degrees of dynamometer tension (N = 60).

Figs. 2 and 3 if we ignore the records obtained under conditions of relaxation and under maximum tension.

RELIABILITY

In Experiment IV, 30 knee-jerks were recorded for each subject under conditions of no experimentally induced tension in an attempt to provide opportunity for adaptation to occur before the main part of the experiment. Each of the 60 subjects in this experiment returned to the laboratory within 3 days, and at the same hour, for a continuation of the experiment. During the second experimental session, 30 responses were again recorded under no experimentally in-

duced tension. The correlation between the mean amplitudes obtained on these two days was $r = .75$. This suggests that individual differences in the amplitude of response may be reliably measured by this method.

The mean amplitude of the records obtained under conditions of no experimentally induced muscular tension was 14.6 mm. The S.D. of the distribution was 14.6. Under 3/4 dynamometer tension, the mean was 33.5, and the S.D. 18.8. These values indicate the range of individual differences in the amplitude of response.

The correlation between the records obtained under 0 and 3/4 dynamometer tension was found to be $r = .65$, which suggests that this measure of tension may be used in comparing individual differences in the amount of change in muscular tension resulting from various experimental conditions.

It was thought that the correlation between responses obtained under 0 and 3/4 dynamometer tension may have been attenuated as a result of tension induced by the experimental situation itself, without the dynamometer effort. If such were the case, the response under 0 tension would not actually indicate the 'normal' amplitude of response. Furthermore, if individual subjects reacted differently to the experimental situation, the base-line of 'normal' response would not represent a comparable condition of tension for the group. It might be expected, however, that any increase in tension produced by uncontrolled factors in the laboratory situation would be very slight in comparison with that induced by squeezing the dynamometer. Thus, a slight degree of dynamometer tension would serve to obscure any effect due to such uncontrolled factors. This would lead us to predict that responses obtained under a slight amount of dynamometer tension would correlate higher with those obtained under 3/4 tension than would those obtained under the so-called 'normal' condition of 0 tension. This was actually found to be the case. The correlation between responses under 1/8 and 3/4 dynamometer tension was $r = .79$.

The consistency of the measurement of quadriceps-thickening under the various degrees of muscular tension is

indicated by Fig. 5, where the mean amplitude of the odd-numbered responses is compared with the mean amplitude of the even-numbered ones for each degree of tension. The similarity of the curves offers evidence of the consistency with which changes in tension may be measured by this method.

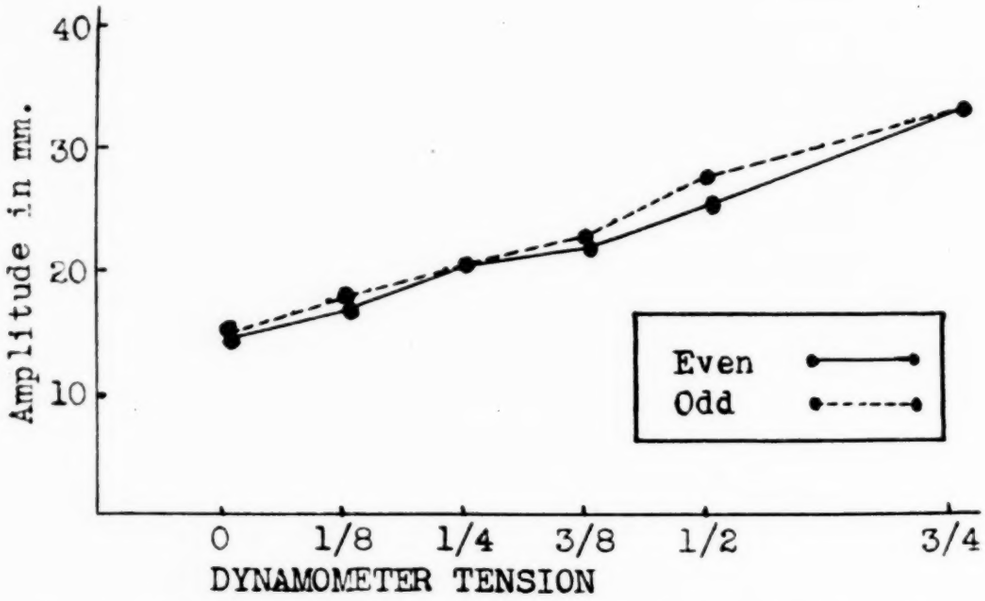


FIG. 5. Mean amplitudes of even and odd numbered responses under various degrees of tension (N = 60).

SUMMARY AND CONCLUSIONS

Eighty-one male undergraduate subjects participated in an experiment in which the knee-jerk was recorded during the exertion of various amounts of muscular effort. The response was recorded mechanically in terms of the thickening of the quadriceps muscle in response to the tendon-tap.

It was found that the amplitude of the knee-jerk bears a linear relationship to the amount of tension exerted on a Smedley hand-dynamometer within the range of no tension to dynamometer tension equal to 3/4 of the reading obtained by the subject at the end of 30 seconds of maximum effort in a preliminary test.

A comparison of means based on two sets of 5 responses for each of 60 subjects suggests that the amplitude of the knee-jerk offers a consistent index of changes in muscular tension.

(Manuscript received October 20, 1938)

References

1. Dodge, R., and Travis, R. C., The relationship between muscle tension and muscle thickening, *Amer. J. Psychol.*, 1930, **42**, 295–297.
2. Golla, F. L., and Antonovitch, S., The relation of muscular tonus and the patellar reflex to mental work, *J. Ment. Sci.*, 1929, **75**, 234–241.
3. Jacobson, E., and Carlson, A., The influence of relaxation on the knee-jerk, *Amer. J. Physiol.*, 1925, **73**, 324–328.
4. Jacobson, E., Differential relaxation during reading, writing, and other activities as tested by the knee-jerk, *Amer. J. Physiol.*, 1928, **86**, 675–693.
5. Jacobson, E., *Progressive relaxation*, Chicago: Univ. of Chicago Press, 1929.
6. Lee, M. A., and Kleitman, N., Studies in the physiology of sleep, II, Attempts to demonstrate functional changes in the nervous system during experimental insomnia, *Amer. J. Physiol.*, 1923, **67**, 141–152.
7. Lombard, W. P., Variations of the normal knee-jerk and their relation to the activity of the central nervous system, *Amer. J. Psychol.*, 1888, **1**, 2–71.
8. Tuttle, W. W., The effect of attention or mental activity on the patellar tendon-reflex, *J. Exper. Psychol.*, 1924, **7**, 401–419.

REPETITIVE PATTERN IN WHOLE AND PART LEARNING THE SPIDER MAZE

BY THOMAS W. COOK

Acadia University

Certain aspects of data derived from studies of response to relationships at implicit and explicit levels make it seem probable that the relative economy of whole and part methods of learning might be affected by the presence of repetitive patterns in the material. As far as I can discover, however, the only experimental attack on this problem is Seagoe's experiment on the significance of qualitative vs. quantitative wholes (4), and the sole cogent theoretical analysis that by Commins (1). However, while I believe that Commins has rightly understood much of the significance of the general work on response to relationships for the whole-part problem, it would be desirable that his conclusions be tested by experiments designed to supplement Seagoe's brilliant beginning in this difficult field. Since the multiple choice nature of the 'spider maze' (2) permits the construction of a sufficiently comprehensive range of patterns, it seemed feasible to use these mazes as material for a further essay at the problem.

Eight each of 3 types of maze pattern were used: part-pattern, whole-pattern, and nonsense mazes. Four mazes of each type were learned by the whole method, and 4 mazes by mastering each of the four parts separately and afterwards tracing the maze as a whole. The nonsense mazes were similar to those used in earlier experiments (2, 3) and served as controls. A part-pattern maze contained 4 repetitive patterns, one for each 8-unit part. Thus the first part-pattern maze may be represented by the numbers 42322212, 12344321, 44332211, 13311331. These numbers signify the number of staples between entrance and exit on a unit, and follow the system of counting used by subject W (3, p. 440): on the first unit the exit is 4 staples from the entrance, on the second unit 2 staples from the entrance, etc. Since each repetitive pattern was specific to an 8-unit part, and each part had a different pattern, it seemed probable that any identity between successive sections of a maze would be most readily discovered by massing trials on an 8-unit part without intervening tracing of other unrelated portions of the maze: that is, by the part method. Moreover, subsequent integration of the parts into a whole would offer little difficulty once the patterns had been discovered,

since the pattern of each part could be apprehended as a unit, and the four parts would be within the 'span of apprehension.'

The whole-pattern mazes, on the other hand, were constructed with a view to penalizing the part method and favoring the whole method. Each maze was composed of a simple pattern of 5 numbers repeated six and two-fifths times. Since the 'parts' in part learning were of 8 units, the end of a part did not coincide with the end of any 5-unit pattern, and if the 4 parts of a maze were compared unit for unit correspondence between units would be accidental and infrequent. Thus the parts of maze 2 (pattern 32414) were 32414324, 14324143, 24143241, 432414321. It seemed possible that the emphasis placed by part procedure upon the 8-unit sections would make the discovery of 5-unit patterns more difficult than in whole learning, where the situation does not favor this particular 'set.'

Besides the whole and part-patterns, a further repetitive feature was introduced. The 'true path' around the central block alternated from unit to unit RLRLRL or LRLRLR for all mazes used in the present experiment. According to maze type the order of learning was part-pattern, whole-pattern, nonsense, nonsense, whole-pattern, part-pattern: repeated 4 times. According to method of procedure the order of learning was (P—part, W—whole): PPPWWWPPPWWWWWPPPPPPWWW for subject W, WWWPPPWWWPPPPPPWWWWWPPP for subject M, and WPWPWPP-WPWPWWWWPPPPPPWWW for subject C. One additional minor change was made in procedure. Only one errorless trial was required in combining the parts or in whole learning. Three errorless trials remained the standard for initial part learning. The first 6 mazes were learned before and the remainder after the Xmas recess, 1937. With this exception the time interval between successively learned mazes was from 1 to 6 days. Otherwise the procedure was identical with that of earlier experiments with the spider maze (**2, 3**).

Experimental Data

I

Nonsense Mazes

Since the nonsense mazes differed from the patterns used in our previous work with 32-unit spider mazes only in specific details and in the alternating character of the right-left choices, it is not surprising to find similar results.

A. First Trial

As in the earlier work, first trial errors show no significant differences between the relative economy of part and whole methods, no practice effects from one maze to the next, relatively small regional and individual differences, and difficulty of tracing directly proportional to size of pattern. The average number of errors per maze is about 7 percent less in the present experiment, probably due to the alternation of

R–L choices around the central block. For time similar considerations hold. Subject W (as before) requires more time for first trial tracing with the part method, but in the present investigation the differences are not large enough to be reliable.

B. Later Trials

1. *Whole vs. Part Procedure*

As Table I indicates, the results for later trials with the nonsense mazes are identical in general trend with the findings of the previous experiment (**3**). For all three subjects the part method is markedly superior to the whole method in errors and moderately superior in time, with little or no difference in number of trials required by either method. Since only 4 nonsense mazes were learned by the part method and 4 by the whole method, we cannot attach too much importance to the data from a single subject, but the averages for all 3 show about 3.3 times as many errors and 28 percent more time by the whole method. These values are somewhat more favorable to part procedure than was the case with the earlier work, indicating that practice from one maze to the next is still favoring the part method.

TABLE I

Maze Type	Method of Learning	Errors			Time*			Trials		
		Subject								
		W	M	C	W	M	C	W	M	C
Nonsense	Whole	185	141	205	1159	916	1051	9.0	7.2	7.0
	Part	41	61	61	721	757	966	7.5	7.3	7.5
Whole-pattern	Whole	19	178	274	375	1046	1248	3.7	7.7	8.2
	Part	21	57	122	538	798	1467	6.0	7.9	10.2
Part-pattern	Whole	73	157	182	537	856	1007	5.5	6.5	7.0
	Part	13	35	90	394	578	1138	4.8	5.7	8.3

* Time in seconds.

2. *Amount of Material*

(*a*) *Relative Difficulty of Parts I, II, III, IV.*—Regional differences in initial part learning are slight. Since, as we

shall see later, C and M (except in two instances) learned all mazes as if they were nonsense materials, it is in order for present purposes to pool their results for all three types of maze. These, together with the nonsense mazes for subject W, yield averages (to the nearest whole number) of 12, 11, 13, and 14 errors for parts I, II, III, and IV respectively.

(*b*) *Relation of Size of Pattern to the Whole-Part Problem.*— The ratio of errors in 8-unit parts to 32-unit wholes is slightly less with C and M, and slightly greater with W than in the earlier experiment. The increased superiority of the part method in the present investigation comes wholly from enhanced ability to combine the parts, particularly with subject C. The average errors for subjects W, M, and C respectively are 27, 6.5, and 10.5 per maze, or 66, 11, and 17 percent of the part method totals.

II

Part-Pattern and Whole-Pattern Mazes, Subjects M and C (Omitting the Last Two Mazes Learned by Subject M with the Part Method)

From M and C, first trial records with part-pattern and whole-pattern mazes are nearly identical with those for the nonsense mazes. With later trials, except for the last two mazes that M learned by the part method, C and M found the patterned structure of the part-pattern and whole-pattern mazes no assistance in learning. In fact (with the exception noted) the only quantitative indication of response to pattern was that C made poorer records on the whole-pattern mazes than on either of the other two types. Subject M also did not do so well on the whole-pattern mazes, and since this result appears for both part and whole methods and with all 3 measures, it is possible that the repetition of a 5-unit pattern throughout a maze was sufficiently sensed to give rise to some 'associative inhibition.'

The relative economy of whole and part procedure, and the relations between the factors conditioning that economy, are also similar to the results for all 3 subjects with the nonsense mazes. The only difference is that with the patterned

mazes the part method is slightly less efficient relative to the whole method. In the whole-pattern and the first 4 part-pattern mazes M made 3 times as many errors with the whole as compared with the part method, but C made only slightly over twice as many errors in whole as in part procedure. With time and trials the averages for the two subjects show no advantage for either method.

In relation between size and difficulty, and (with one exception) in combining the parts, the results also parallel those for the nonsense mazes. The exception is that subject C had considerable difficulty in integrating the whole-pattern mazes: her error average was 50 errors per maze, compared with 10 errors for the nonsense mazes and 12 errors for the part-pattern mazes.

III

Part-Pattern and Whole-Pattern Mazes, Subject W; Last Two Part-Pattern Mazes Learned by the Part Method, Subject M

The results for subject W with the part-pattern and whole-pattern mazes, and for subject M with the last two part-pattern mazes learned by the part method, stand in marked contrast to the remainder of our data. The averages given in Table I indicate that subject W learned the 'patterned' mazes much more easily than the 'unpatterned,' and that the gross results seem to favor our prediction: part-patterns favor the part method and whole-patterns the whole method. For the part-pattern mazes W found the part method superior in errors, time, and trials, 450, 40, and 15 percent, while with the whole-pattern mazes the superiority of the whole method with these three measures is 10, 44, and 120 percent.

A more detailed analysis of the data, however, clarifies the nature and causes of these phenomena. For the whole-pattern mazes (in trials other than the first) subject W made 46, 14, 8, and 16 errors in the four mazes learned in that order by the part method, and 25, 16, 14, and 23 errors with the mazes learned by the whole method. It would seem, therefore, that if we except the first maze learned by part procedure,

the whole method is, if anything, inferior. Since the first whole-pattern maze that W learned by the part method was number 2, and the first whole-pattern maze learned by the whole method was number 7, the slight advantage of the whole method in average errors may be due to the obvious practice effect (from 46 errors with maze 2 to 25 errors with maze 7). For the part-pattern mazes, however, the advantage of the part method seems genuine. But subject W's results suggest that it is not an absolute superiority of part over whole procedure, but is specific to a certain level of attainment. For the part-pattern mazes W made 45, 4, 2, 3 errors by the part method with mazes 1, 7, 18, 19; and 101, 150, 35, and 8 errors by the whole method with mazes 6, 11, 13, and 24. It would seem clear that during or immediately after working with the first part-pattern, W mastered an improved technique for learning the part-pattern mazes by the part method, but did not reach a similar achievement with the whole method until the last maze learned. Subject M's results are similar though not so striking. He made 178, 133, 202, and 115 errors with part-pattern mazes learned in that order by the whole method, and 46, 52, 32, and 10 errors for part-pattern mazes learned by the part method.

Examination of W's method of learning and of her spontaneous comments throws more light on the matter. I have described previously the manner in which W learned by counting aloud and thereby discovering and verbalizing a number corresponding to the number of staples from entrance to exit on each unit, and subsequently memorizing the serial order of these (latter) numbers (**3**, p. 440). In the present instance this method proved of great assistance: rehearsing the numbers aloud seemed to call attention to the pattern. On maze 1 she made no remarks except to say that the parts were easy, but the directions around the central blocks confusing (It was a part-pattern maze). Halfway through the first trial with maze 2, the first whole-pattern maze, she said: "There's some pattern, but I can't tell what." She made 89 errors on that trial. As usual for all types of maze, W had little difficulty in subsequent trials with the separate parts

making 2, 0, 0, and 3 errors in parts I, II, III, and IV. But on the first attempt to combine the parts she made 40 errors. At the end of that trial she remarked: "I can't get a pattern to stand out, but I feel that it's just the same thing over and over." (The pattern was 32414, 32414, etc.) But on the second attempt she made only one error, and that a directional one.[1]

On number 6, the next maze of the part-pattern type (learned in this instance by the whole method), W made no comment regarding pattern, although she showed signs of responding to some formal features of the first 3 'parts.' But the total of 104 errors shows no marked increase in economy of learning. At this point, however, W discovered the alternating character of the pattern of 'correct' choices of direction around the central block from unit to unit. At the end of trial 4 she said, "You go one side one time, and the other side the other time, right straight through." Hereafter she made use of the discovery, often remarking on seeming deviations from it. However, she never noticed that the alternating pattern only held if the blocks on which the exit was 180° from the entrance were omitted. I had in fact anticipated a possible difficulty with this point, and had to guess whether including or omitting the 180° blocks as part of the alternating pattern would be more confusing. Apparently I guessed wrongly.

On the second whole-pattern maze (number 5, the first one learned by the whole method) W remarked at unit 20: "There must be something to this." But it was not until unit 30 that the 'something' became clear, and she made no errors on units 31 and 32. In trial 2 the third unit gave some trouble, but thereafter all errors were directional. From that time W was on the alert for whole-pattern mazes, and never failed to discover their pattern somewhere between units 8 and 12 on the first trial. The signs of emergence of pattern were unmistakable. Until it appeared W would go cautiously, count aloud, and watch closely for repetition of numbers. And as soon as

[1] The correct number of staples from the entrance, but in the wrong direction (R or L) around the central block on a given unit.

she thought she had found a pattern she would start applying it. This meant a duality of method in first trial tracing: the method used in previous work until the pattern was found, and the application of that during the remainder of the trial. One result of this procedure was to reduce the number of first trial errors. In fact W made only an occasional (directional) error after discovering a given pattern. But the patterns emerged with equal ease in part and whole procedure. The average number of first trial errors on whole-pattern mazes was 43 for the part method and 41 for the whole method.

For the part-pattern mazes learned by the part method the total number of first trial errors was 86, 58, 58, and 39 in order of learning. Moreover, if we omit the first and the initial quarter of the second of these mazes, which show no evidence of grasp of pattern on the first trial, the first 5 units of the 8-unit parts contain almost all the errors: 116 errors in the first 5 units with a chance expectation of 165 (70 percent), as against 16 errors in the last 3 units with a chance expectation of 99 (16 percent). Since each 8-unit part of the part-pattern mazes had a specific pattern which a subject could only discover while tracing that section, the discovery of pattern could only show during trial 1 at the point where it did appear —in the last few units of each part. For later trials the three part-pattern mazes last learned show only directional errors.

With the whole method the only 2 part-pattern mazes to give evidence of grasp of pattern by increased efficiency were maze 13 (35 errors after trial 1) and maze 24 (10 errors after trial 1). In the former W made 80 first trial errors, slightly above her average for the nonsense mazes, but in the latter only 40 errors in trial 1. Moreover, 33 of these errors occurred in the first 5 units of each part, with only 7 errors in the last 3 units.

Subject M, though his quantitative records give no indication of the fact, showed signs of responding to pattern with maze 12, a part-pattern maze learned by the part method. In parts I and II he reported nothing, either while tracing or in the spontaneous retrospection to which he was addicted whenever anything about a maze struck him as unusual. But at

the end of the first trial on part III he said: "My, that's simple, 2200, 2200."[2] On part IV he also found the pattern during the first trial, although as in part III he made a few later trial errors. Maze 13, another part-pattern maze, was also learned by M with part procedure. Here he grasped the structure of the first two parts on trial 1, and announced each time "I know your maze." But at the end of part III he said: "No pattern to this one," and with part IV also failed to find any repetitive configuration. Here the later trials corroborate the preliminary analysis. He made 5, 1, 14, and 10 later trial errors in parts I, II, III, and IV respectively. For the last part-pattern maze that M learned by the part method, number 24, he offered only one comment (at the end of trial 1, part IV), "Now, I know what that is if I could only figure it out." His first trial errors were about the same as previously found for the nonsense mazes, but in later trials he made only 6, 4, 0, and 0 errors in parts I, II, III, and IV, and no errors in combining the parts. With the whole method the only evidence of response to pattern was a remark made while tracing part-pattern maze number 18: "That would be easy by the part method."

Discussion

That subjects may fail to find repetitive patterns, and, on the other hand, that the number system may be of great benefit in discovering, holding, and subsequently using such patterns, are not altogether novel conclusions, though to date they have been little demonstrated in maze experiments, and have seldom been so conclusively shown for any type of material. The interest of the experimenter, however, is more apt to be directed toward the gradations of response intermediate between no grasp of pattern on the one hand and explicit formulation on the other. In this territory lie facts of the greatest significance for the understanding of problem-solving. Probably because of the complexity of our patterns, our set-up did not lend itself to implicit forms of response to

[2] W would have said 3311, 3311, since she counted the exit on each unit while M did not.

relationships, and our contributions to the problem are accordingly meagre.

First, we note that C learned the whole-pattern mazes less efficiently than the nonsense mazes. Taken at its face value, this might mean that a rudimentary grasp of pattern led to interference between successive sections of a maze. Other than this C showed no indication of response to any but one or two of the simplest part-patterns. Thus with the pattern 12121212 C seemed to sense the identity of the four 1's, but showed no insight into the likeness between the 2's. I am inclined to attribute C's failure in this connection to certain features of her method of tracing which made the use of numbers difficult if not impossible. For example, I have records of hundreds of instances in which C changed the direction taken around the central block. And these instances occur at any and every stage in the learning process. Sometimes when the exit from a unit was the staple adjacent (say to the right) of the entrance, C would take the shorter path for 6 or 8 trials, and on the next trial *go the long way around* without hesitation or entrance into one of the 6 intervening culs-de-sac. And, when encouraged to talk, she would give no indication that she had traced the unit in a manner different from preceding trials. Obviously C could not be relying upon numbers, since to get the correct count for the longer route one must subtract 1 from 8, and then count 7 staples. Such a proceeding occurring for the first time in a series of trials could scarcely escape her notice.

W used numbers throughout, and in such a fashion as to exclude more implicit forms of response to relationships. M also seemed to use numbers whenever he got into difficulties, though to what degree I cannot be certain, since I purposively avoided questions which might suggest to the subjects the presence of repetitive patterns. It seems probable, therefore, that the spontaneous statements of W and M regarding the presence of patterns which they were as yet unable to grasp clearly refer to the 'feel' of a pattern within the numbers rather than some more implicit sort of postural or imaginal data. I am indeed inclined to suspect that discovery of

complex repetitive maze patterns is impossible without the use of numbers.

Neither would I be inclined to overestimate the breadth of application of our results regarding the relation between discovery of pattern and the whole-part problem. As far as they go, our data indicate that the part-pattern mazes favor part procedure, but show little or no difference in relative economy of the two methods with the whole-pattern mazes. Also, the difference in relative economy of the whole and part procedures with the part-pattern mazes seems to concern ease of discovering the correct mode of attack, and it is accordingly quite possible that subjects highly trained in the application of the number system to this type of problem would be equally efficient in part and whole procedures. However, it must be remembered that we have sampled only a few of the infinite varieties of type and arrangement of pattern. For example, it is probable that once subject W had become familiar with 5-unit patterns the whole-pattern mazes became too easy to differentiate between the two methods. More complex patterns and larger mazes might give different results.

Summary

The significance for the whole-part problem of identity between successive sections of 32-unit spider mazes was investigated with 3 practiced subjects learning 8 mazes each of 3 types: (*a*) a 5-unit pattern repeated throughout, (*b*) 2-unit or 4-unit patterns each specific to an 8-unit part, (*c*) nonsense mazes containing no repetitive patterns. Apart from the nonsense mazes, which give results almost identical with previous work (**2, 3**), the chief findings of the experiment were:

1. The presence of repetitive patterns in maze material is no guarantee that subjects will respond to them.

2. Discovery of repetitive pattern may lead to a large and sudden increase in the efficiency of tracing.

3. The use of numbers to signify the relation between entrance and exit within a unit and the sequence of maze units favors the emergence of repetitive pattern.

4. Within the limits of the experiment, 8-unit part learning

of 32-unit mazes favors discovery of patterns specific to those 8-unit parts, but 5-unit patterns repeated throughout 32-unit mazes are learned with equal economy by whole and (8-unit) part procedure.

(Manuscript received October 29, 1938)

References

1. Commins, W. D., *Principles of educational psychology*, Ronald Press, New York, 1937, 409–411.
2. Cook, T. W., Whole versus part learning the spider maze, *J. Exper. Psychol.*, 1937, **20**, 477–479.
3. Cook, T. W., Whole and four part learning thirty-two unit spider mazes, *J. Exper. Psychol.*, 1938, **22**, 439–450.
4. Seagoe, M. V., Additional laboratory experiments with qualitative wholes, *J. Exper. Psychol.*, 1937, **20**, 155–168.

A NOTE ON LABORATORY STUDIES OF DECEPTION

BY F. K. BERRIEN

Colgate University

Laboratory experiments which have sought to validate various indices of deception have been thwarted at times by the difficulties involved in setting up satisfactory 'laboratory crimes.' The purpose of this communication is to suggest a possible convenient solution to the problem. Most attempts to create laboratory crimes are open to one or more of the following criticisms.

In the first place, the manufactured situations are frequently quite unreal, thereby failing to place the subject in a situation comparable to a genuine suspect in an actual crime. It is obvious to everyone that the subject has little at stake. Within the laboratory the subject's life, liberty, reputation or property do not depend upon his ability to concoct a satisfactory alibi.

Some investigators have attempted to solve this problem by betting with the subject as to whether or not the experimenter can detect falsehoods on the basis of the polygrams. This method has the merit of increasing the emotional tension but sometimes conflicts with the wishes of college administrators. At best the procedure still smacks of artificiality.

In the second place, the effectiveness of the laboratory crime in producing the desired emotional responses on the part of the subject often depends upon the ingenuity of the experimenter. Runkel [1] required his subjects to follow a set of instructions which led them through what appears to have been a very exciting series of events. Burtt,[2] by contrast, had

[1] Runkel, J. E., Luria's motor and word association in the study of deception, *J. Gen. Psychol.*, 1936, **15**, 23–37.

[2] Burtt, H. E., Motor concomitants of association reaction, *J. Exper. Psychol.*, 1936, **19**, 51–63.

his subjects merely examine the contents of either one of two boxes. One of these contained objects suggesting a struggle over a card game such as a bullet, a deck of cards with five aces, and a knife with blood on it. The other box contained feminine articles such as a powder puff, some hair pins, and cosmetics. Earlier, Burtt [3] devised a procedure whereby the subject was provided with a set of facts which supposedly had been established by sources other than the suspect. If during the experimental trial the subject chose to tell the truth he was provided with an alibi written by the experimenter's assistant. If the subject preferred to lie he was compelled to invent his own alibi and use it as if he were attempting deception. Note, however, that whether the subject lied or not he was called upon, in effect, to act a part.

It must be apparent that the experimental procedures employed in these studies vary over a wide range as exemplified by the cases cited above. It is difficult to imagine that they are all equally successful in producing emotional reactions that are comparable to those that are incident to 'real lying.' We seem to be led to the very peculiar conclusion that the validity claimed for a particular index of deception in part at least reflects the experimenter's cleverness in arranging the experimental details.

In the third place, a further problem is thereby introduced. If the above conditions are not equally successful in evoking emotional responses, then comparisons of emotional indices on the basis of validity are meaningless unless these indices are recorded simultaneously or unless the experimental conditions are equivalent when the indices are recorded. The latter assumption is often made but frequently without sufficient foundation.

These three difficulties seem to have their roots in something more fundamental. There has been little effort to make clear the precise psychological reactions that take place when an individual is forced to lie. It may well be that the obvious difficulties mentioned above can be overcome by a procedure

[3] Burtt, H. E., Inspiration-expiration ratio during truth and falsehood, *J. Exper. Psychol.*, 1921, 4, 1–23.

that suggests itself when these reactions are listed and examined.

If one stops to reflect upon the temporal sequence of reactions that take place in lying, three rather distinct periods emerge. There is first of all the initial startle when a critical question is asked. This is followed by a tension period of varying length during which an explanation or alibi is being formed. Finally comes a period when the subject rethinks his answer, fearful that the falsehood may be discovered.

These considerations permit us to see clearly the requirements that must be met if we are to approximate the emotional experiences of deception. It is proposed that stimuli of quite a different nature might produce essentially the above pattern of emotional responses even though the thought content might be quite different.

In a group of experiments to be reported in detail later the subjects were given a series of mental arithmetic problems to solve. A pair of binding posts was mounted on an adjustable stand so they could be brought near the subject's face. Across these posts was stretched a piece of fuse wire which when 'shorted' gave a small explosion. Out of the small cloud of smoke came a shower of sparks which might fall on the subject's hand or arm. The subject was protected, however, from serious harm by an intervening piece of glass.

The subject received the following instructions:

> I am going to ask you to do some mental arithmetic and I want you to get the answers as quickly as possible. If you take too long to solve the problems this fuse will blow in your face like this (demonstrate). If some of those sparks should fall on your arm they might give you a slight burn. Now, once you give me an answer don't attempt to correct it. If it is wrong, the fuse *may* blow about 10 or 15 seconds after you give the answer. You see, the fuse may blow anytime after I dictate a problem except when you get the *correct* answer *quickly*.

By this arrangement it was thought the essential lying pattern of emotional reactions would be evoked. The situa-

tion involves the initial period of excitement when the question is asked. The subjects think, "Can I get this one? Will I make a fool of myself on a simple problem this time?" The suspense and anticipation immediately preceding a mental arithmetic problem appears to be similar to that which precedes attempted deception. The period during which the answer to the problem is being worked out is similar in nature if not duration to that in the lying situation. This period in both cases seems to be characterized by an intellectual search for an acceptable answer. Possible answers are checked, inhibited, rechecked to make sure they will "pass muster." Moreover, the whole process must be done under the pressure of time. In the lying situation the subject feels he must not delay for fear of throwing suspicion on himself. In our laboratory situation he must get the answer before that fuse blows in his face.

Even after the answer is given there is the possibility that it is not correct and the fuse may yet blow. He may realize he has given the wrong answer and may sit waiting for the explosion which may never come. This last period of suspense appears to be in all essential respects similar to the lying situation where the subject wonders whether he will be caught in his own story. He is not sure that the alibi will convince his examiner. On the other hand, if the individual gives the correct answer quickly and is positive of its correctness he is not apprehensive concerning the fuse. This situation corresponds to the subject's telling the truth, in which case his "conscience is clear" and he has no fear of what may come.

This arrangement of the stimulus conditions rather effectively meets the difficulties mentioned at the beginning of this paper. There is nothing artificial about the situation. The threat of punishment is real and effective. Faradic shock might be used instead of the arrangement described. Another possible form of punishment might be found in an arrangement whereby a small hammer fitted with pins could be dropped on the subject's arm or hand. The essential requirement seems to be an arrangement with a sword-of-Damocles aspect.

In addition to presenting a situation which is far from

artificial the use of mental arithmetic in this way enables the experimenter to reproduce essentially the same experimental conditions on numerous occasions. The only limit is the number of problems which the subject finds difficult. Other experimental procedures are usually such that once a series of questions has been used on a particular subject, that subject cannot be employed again unless a new 'crime' is invented. The new situation not only entails the difficulties of composition but may also introduce other variables. The suggestion made herein is not open to that objection. Furthermore, it is a convenient procedure which does not depend upon the ingenuity of the experimenter in inventing new crime situations.

One difficulty has been foreseen in respect to the interpretation of results derived under these conditions. The armchair analysis provides reason for believing the emotional responses to the arithmetic problems may be equivalent to those accompanying falsehoods. Armchair analysis however cannot be accepted as final verification of the facts. Such an analysis is merely the first step. The relationship between the two sets of emotional experiences could perhaps be determined through experimentation.

In the experiments to which I have already referred indirect evidence can be found supporting the validity of our analysis. Subjects were in general more emotionally disturbed when they gave incorrect answers than when their answers were correct. Also, the proportion of incorrect answers accompanied by discernible emotional responses was somewhat greater than the proportion of falsehoods accompanied by discernible emotional responses. Emotional reactions were never detected when truthful answers were given and only rarely when correct solutions to the arithmetic problems were given quickly.

It is hoped that other investigators in the field of emotions and deception may find the procedure helpful.

(Manuscript received November 12, 1938)

A NEW CHRONOSCOPE

BY K. L. HERTEL AND R. E. DUNFORD
University of Tennessee

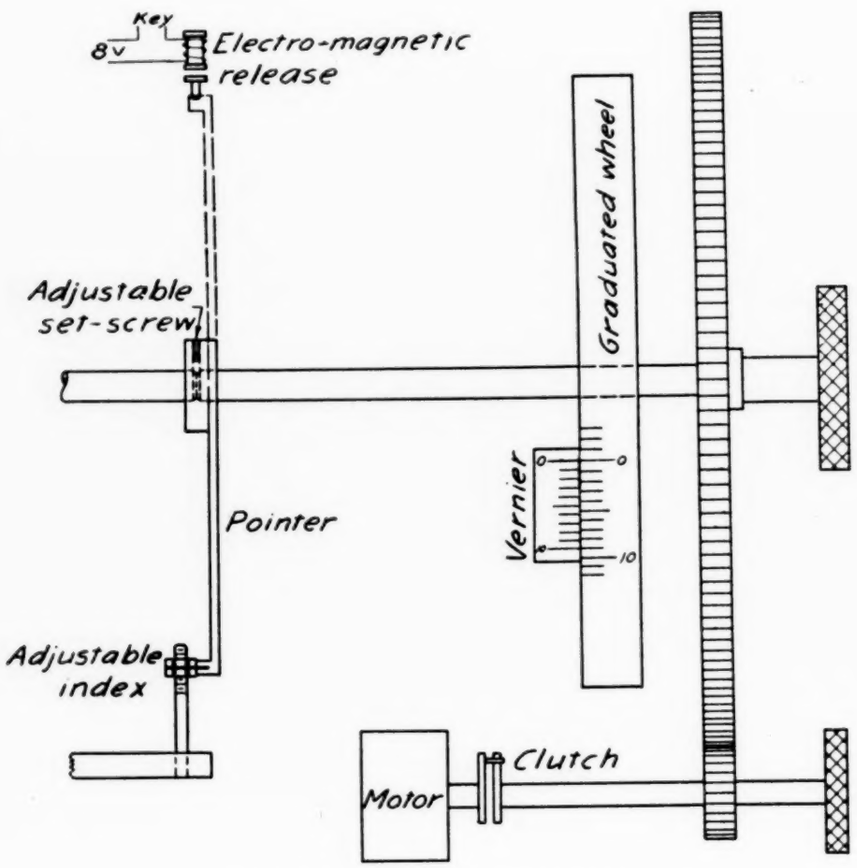
The number and variety of time measuring devices developed or adapted for use in the psychological laboratory are such as to indicate the general interest in such instruments. While general interest in the measurement of Reaction Time as such may have declined in recent years, there is reason to believe that more detailed and analytical studies of this simple behavior may be profitable.

The present chronoscope was designed with a view to eliminating the more serious objections to previous instruments. At the same time it was hoped to add to the instrument certain aspects of accuracy, desirable operational qualities, experimental flexibility, and ease of checking not found in its predecessors.

Convinced that the synchronous motor as used in the Dunlap Chronoscope provided the most acceptable power source, we have used an air-cooled self-starting synchronous motor driven by the usual 60 cycle A.C. current. Through a simple gear the synchronous motor drives a steel shaft on which there are ten indexed pointers and a wheel. The shaft, wheel, and pointers make one complete revolution in four seconds. The wheel is graduated into 400 units providing reading in .01 seconds. A vernier scale on the graduated wheel extends the readings into milli-seconds or thousandths.

The ten pointers can rotate with the shaft by friction or can be stopped by catches. The catches are released by electro-magnets connected with numbered jacks on the front of the chronoscope. Terminal connections for a power source of 8 volts for the magnets are provided on the left end of the instrument. In operation a Central Scientific A.C.-D.C. rectifier is used as a source of power for the magnets. The wheel is stopped in rotation by a manual control, a small

knurled knob on the right end of the chronoscope. All controls, for convenient operation, are located on the right end. These controls are a switch for starting the synchronous motor and setting the mechanism in action; a large knob for manual rotation of the shaft to effect coincidence of indexes when time intervals are to be read; a smaller knob below the larger one permits finer adjustment of indexes; a small lever in the rear of the other controls resets all catches and thus

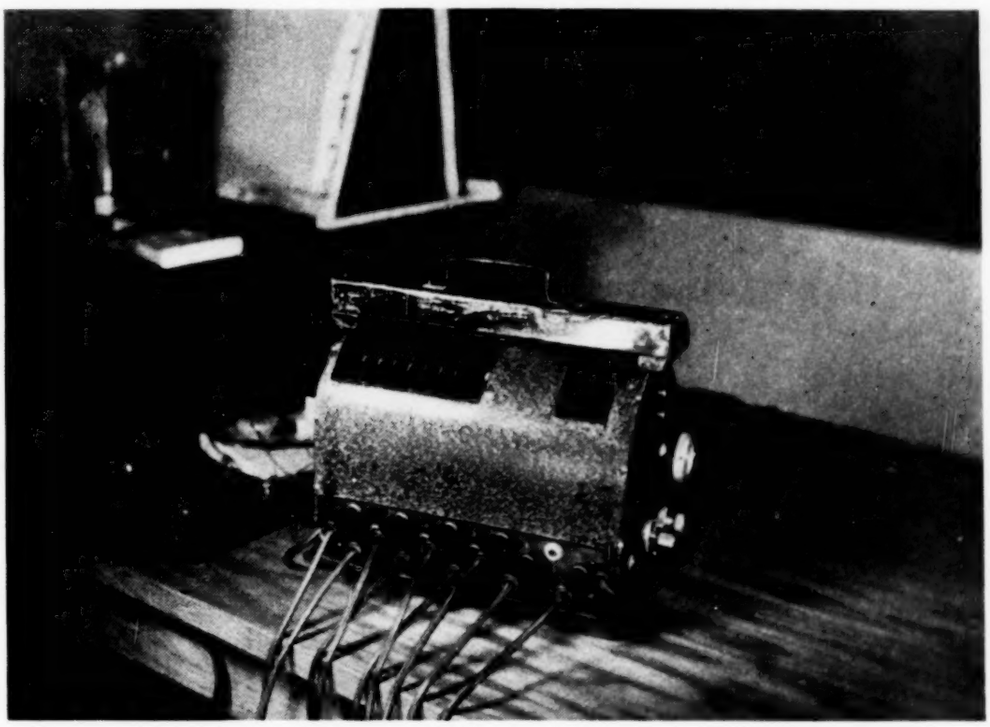


Essential parts of the chronoscope.

provides an automatic setting of all pointers on zero between readings. The entire mechanism is covered with a dust-proof case, with glass observation windows over the graduated wheel and pointer indexes.

In operation the motor turns the shaft on which all ten pointers and wheel are at rest. When released at the incidence of the experimenter or by the subjects, the pointers rotate with the shaft. By friction the pointers are tending

to rotate with the shaft at all times. The construction is such that the lag usually found in magnetic-mechanical clutches is small. The index for each pointer is adjusted to compensate for the pointer-release lag. As long as the lag for the pointer-release is constant no error is introduced. The variation in the lag has been found to be negligible. Furthermore, the pointers are not stopped by any similar device. Tests have demonstrated that no evidence of initial inertia is present in the pointer rotation. The friction of the pointers on the shaft is accomplished by a spring under set screw pressure in



Chronoscope connected for eight subjects.

the collar of the pointer on the shaft. Test runs at high speeds over several hours have shown no appreciable wear or alteration in the coefficient of friction in this arrangement. In experimentation or demonstration the necessary connections with the chronoscope are made simply by plugging the electrically connected keys into the desired number of jacks. Simultaneous reaction times by nine subjects can be taken at one time. This is merely the limitation of the present instrument. It could be extended to any desired number.

To obtain readings the motor is started and the shaft begins to rotate. All pointers are stationary and the graduated wheel is at zero. The presentation of the stimulus releases the experimenter's pointer which rotates with the shaft. The subsequent reaction of the subject releases his pointer. The relative position of these pointers on the shaft is then converted into time by disengaging the motor and rotating the shaft manually until E's pointer coincides with its index. The graduations on the wheel read zero. The wheel is then released and the shaft rotated to coincidence of the subjects's pointer with its index. The time is read directly on the graduated wheel in milli-seconds.

The outstanding merits of the chronoscope are as follows:

1. *Sturdy Construction.*—The chronoscope has been made exceptionally strong and rugged and in this respect is somewhat different from traditional instruments of this type.

2. The instrument can be *quickly tested.* All pointers may be simultaneously released and, if they are not coincidental, the error is immediately revealed.

3. *Simplicity.*—The simplicity of the design has been vindicated in preliminary tests. The elimination of make-break-magnetic-mechanical devices contributing to the tempermental behavior of other chronoscopes has been removed. These tests have indicated that many of the accessory devices such as keys and other electro-mechanical gadgets used in the operation and testing of chronoscopes have contributed much to their error. This chronoscope reveals that it is almost impossible to adjust two keys as to tension and distance between points to produce simultaneous electrical impulse.

4. *Flexibility.*—The possibilities of experimentation by taking a number of reaction times simultaneously and successively need no elaboration. Group experiments, choice, reaction, serial responses, complication problems, etc., are susceptible to timing with this instrument. The advantage of direct reading for each subject is obvious.

(Manuscript received November 9, 1938)

THE PARRISH LABORATORIES OF PSYCHOLOGY AT RANDOLPH–MACON WOMAN'S COLLEGE

BY HELEN PEAK

Randolph-Macon's laboratory of Psychology was established in 1894, one year after the College's opening, with Miss Celestia S. Parrish as the first Professor of Psychology. The experimental tradition which Miss Parrish established here at a time when laboratories of Psychology were rare, has been continued through the years and, in the Fall of 1938, the Department moved into a new laboratory. Two floors of a

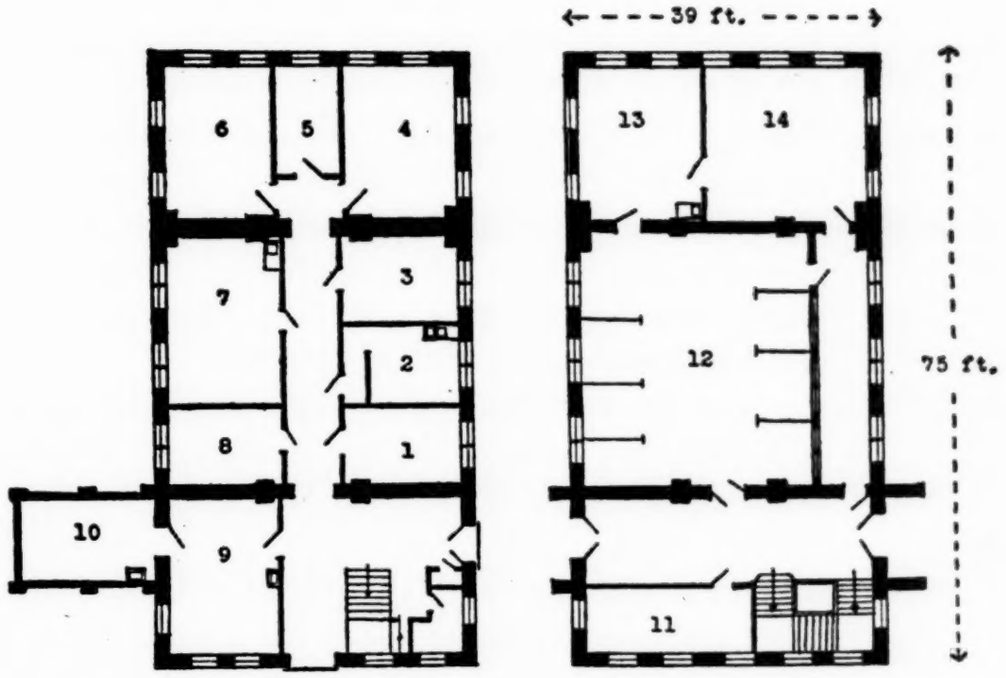


FIG. 1

completely remodelled building provide approximately 5600 sq. ft. of space. The floor plan of the laboratory is shown in Fig. 1, and a description of the rooms is given in the following paragraphs.

Room 1—Staff office. Dimensions, 14′ 6″ × 10′ 6″.

Room 2—Dark room. This room is arranged for visual experimentation as well as for photographic developing. The equipment consists of a double sink and drainboard, safe lights, zinc covered tables, and storage space. A ventilating fan is provided. Dimensions, 14′ 6″ × 9′ 6″.

Room 3—Sound shielded room. This room is specially insulated in order to reduce external and reflected sounds. The equipment consists of a photochronograph, Point-o-lite, and accessory equipment for photographing lid reactions and reactions of other isolated muscle groups. Ventilation is provided. Dimensions, 14′ 6″ × 10′ 6″.

Room 4—Seminar room. The room is equipped with seminar table and chairs. Dimensions, 19′ × 13′6″.

Room 5—Secretary and Assistant's office. Dimensions, 14′ × 8′.

Room 6—Staff office. Equipment for storage of reprints and file cases is provided. Dimensions, 14′ × 19′.

Room 7—Advanced experimental laboratory. This is a combination class room and research room, fitted with tables, chairs, apparatus cases, running water, and gas. Dimensions, 20′ 6″ × 14′ 6″.

Room 8—Research room. The equipment consists of tables and cases for mental testing material. Dimensions, 14′ 6″ × 10′ 6″.

Room 9—Animal room. This room serves to house small animals and is also used for maze experiments with these animals. Storage space for animal food, a sink for washing cages, and zinc covered tables for the animal cages are provided. Apparatus cases at one end of the room are used for general laboratory equipment. Dimensions, 17′ × 19′.

Room 10—Shop. Small hand tools are provided for departmental use. A work bench, work tables, a sink, and a large amount of storage space are provided in this room. Dimensions, 14′ 6″ × 9′ 6″.

Room 11—Departmental library. This small room serves to house the current periodicals and the reference books used by the students in advanced courses in the Department. Shelf space, chairs, and reading tables are provided. Dimensions, 21′ × 8′ 6″.

Rooms 12 and 13—Student laboratory. These two rooms are used for introductory laboratory work, which is provided for all students taking the first course. 6 ft. partitions separate the student work tables in room 12. In room 13 apparatus cases set at right angles to the walls serve as partitions. These two rooms accomodate 28 students at one laboratory session. Equipment consists of 14 small tables, the necessary chairs, a demonstration table for the instructor, apparatus cases, and other storage cabinets. Running water, gas, and electrical outlets are provided. Dimensions, room 12, 32′ × 29′; room 13, 17′ × 15′.

Room 14—Class room for Introductory Psychology. This room is provided with tablet-arm chairs, demonstration table for the instructor, lantern table, and aluminum coated screen. Midnight shades at the windows fit the room for projection purposes. Dimensions, 23′ × 20′ 6″.

Electric outlets are provided at all necessary places on both floors. Special circuits connect rooms 2, 3, 7, and 12, so that interrupted currents for timing or reaction currents from one room can be transmitted to other laboratories. Running water and gas are provided in Rooms 2, 7, 9, 10, and 13. All the walls and floors are sound-insulated.

(Manuscript received February 27, 1939)